BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

202

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A61K 49/00, C12N 5/06, C07H 21/04

A1

(11) International Publication Number: WO 98/51351

(43) International Publication Date: 19 November 1998 (19.11.98)

(21) International Application Number: PCT/US98/10080

(22) International Filing Date: 15 May 1998 (15.05.98)

(30) Priority Data:

08/857,076	15 May 1997 (15.05.97)	US
08/888,534	7 July 1997 (07.07.97)	US

(71) Applicant: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02110–2214 (US).

(72) Inventors: RUVKUN, Gary; 120 Herrick Road, Newton, MA 02138 (US). KIMURA, Koutarou; 9 Hawthorne Place, Boston, MA 02114 (US). PATTERSON, Garth; 106 – 13th Street #321, Charlestown, MA 02129 (US). OGG, Scott; 19 Phillips Lane, Newton, MA 02159 (US). PARADIS, Suzanne; 21 St. James Avenue, Somerville, MA 02178 (US). TISSENBAUM, Heidi; 43 Chandler Street, Belmont, MA 02178 (US). MORRIS, Jason; 564 First Avenue #17Q, New York, NY 10016 (US). KOWEEK, Allison; 9 Belknap Street #3, Somerville, MA 02178 (US). PIERCE, Sarah; 260 Brookline Street #1, Cambridge, MA 02139 (US).

(74) Agent: ELBING, Karen, L.; Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110–2214 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS

(57) Abstract

Disclosed herein are novel genes and methods for the screening of therapeutics useful for treating impaired glucose tolerance conditions, as well as diagnostics and therapeutic compositions for identifying or treating such conditions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS

Background of the Invention

This invention relates to compositions and methods useful for delaying or ameliorating human diseases associated with glucose intolerance. Diabetes is a major disease affecting over 16 million individuals in the United States alone at an annual cost of over 92 billion dollars.

Type I diabetes or insulin-dependent diabetes (IDDD) is an autoimmune disease. In the IDDM patient, the immune system attacks and destroys the insulin-producing beta cells in the pancreas. The central role of insulin in human metabolism is to aid in the transport of glucose into muscle cells and fat cells. The body's inability to produce insulin results in hyperglycemia, ketoacidosis, thirst, and weight loss. In addition, diabetics often suffer from chronic atherosclerosis and kidney and eyesight failure. A patient with IDDM requires daily injections of insulin to survive.

The most common form of diabetes is non-insulin dependent diabetes (NIDDM) or Type II diabetes. Type II diabetes is a heterogenous group of disorders in which hyperglycemia results from both impaired insulin secretory response to glucose and decreased insulin effectiveness (i.e., insulin resistance). Older people who are overweight are at particular risk for Type II diabetes. Genetic studies have suggested that, Type II diabetes is found in families and that the disease may be due to multiple genetic defects. In addition, the link between obesity and Type II diabetes is strong. Approximately 80 percent of Type II diabetics are obese. Weight loss and exercise can be effective to keep blood glucose levels normal, reducing the long-term complications of the disease.

At present there are few reliable methods for presymptomatic diagnosis of a genetic predisposition for diabetes or obesity. The search for genetic markers linked to diabetes and obesity has proven difficult, and this is especially true for Type II diabetes.

Treatments for diabetes emphasize control of blood glucose through blood glucose monitoring. The majority of patients take oral medications and/or insulin injections for appropriate control. Treatment of diabetes is generally chronic and lifelong, and treatments are generally not satisfactory over the long run. In addition, insulin treatment may become increasingly ineffective as the disease progresses. While insulin has been known for decades, and within the past decade, the receptors for insulin and aspects of its signaling pathway have been identified, the transcriptional output from these signaling pathways have not been characterized. In addition, the molecular basis of the obesity-induced insulin resistance is unknown.

Summary of the Invention

We have discovered that the *C. elegans* metabolic regulatory genes *daf-2* and *age-1* encode homologues of the mammalian insulin receptor/PI 3-kinase signaling pathway proteins, respectively. We have also discovered that the DAF-16 forkhead protein represents the major transcriptional output of this insulin signaling pathway. For example, we have discovered that it is the dysregulation of the DAF-16 transcription factor in the absence of insulin signaling that leads to metabolic defects; inactivation of DAF-16 reverses the metabolic defects caused by lack of insulin signaling in *C. elegans*. Finally, we have found that the *C. elegans daf-7*, *daf-1*, *daf-4*, *daf-8*, *daf-14*, and *daf-3* genes encode neuroendocrine/target tissue TGF-β type signal transduction molecules that genetically interact with the insulin signaling pathway. Similarly, we have shown that the metabolic defects caused by lack of

neuroendocrine TGF-β signals can be reversed by inactivation of the DAF-3 transcription factor.

Together, this evidence indicates that the DAF-16, DAF-3, DAF-8, and DAF-14 transcriptional outputs of these converging signaling pathways regulate metabolism. In addition, these discoveries also indicate that insulin and TGF-β-like signals are integrated in humans to regulate metabolism, and that the homologues of other DAF proteins are likely to be defective or down regulated in human diabetic pedigrees as well as obesity induced diabetes. These results therefore indicate that the *C. elegans daf* genes are excellent candidate genes and proteins for human disease associated with glucose intolerance, e.g., diabetes, obesity, and atherosclerosis. Our findings indicate that the human homologues of these *daf* genes and proteins mediate insulin signaling in normal people and may be defective or mis-regulated in diabetics. Moreover, our findings indicate that there are at least two classes of type II diabetics: those with defects in the TGF-β signaling genes, and those with defects in insulin signaling genes. Below we describe exemplary sequence and functional characteristics of the human homologues of the daf genes.

The discovery of converging DAF-7 and DAF-2 insulin-like signaling indicates that many cases of obesity-induced and genetically-induced diabetes (and obesity) may be treated by administration of a human DAF-7 polypeptide.

The discovery that defects in the TGF-β signaling pathway can be suppressed by decreases in DAF-3 pathway activity and that defects in the insulin pathway can be suppressed by decreases in DAF-16 activity highlight the utility of transcriptional regulatory DAF proteins in drug development; in particular, drugs that inhibit the activity of these proteins are useful for reversing the effects of obesity-induced or genetically-induced defects in DAF-7 TGF-β type or insulin signaling.

In one aspect, the invention features a substantially pure preparation of a

DAF-2 polypeptide, which can be derived from an animal (for example, a mammal, such as a human, or an invertebrate, such as *C. elegans*). In preferred embodiments, the DAF-2 polypeptide has insulin receptor (InR) activity, insulin receptor related activity, insulin-like growth factor receptor (IGF-1) receptor activity, or a combination of these activities.

The invention also features isolated DNA encoding a DAF-2 polypeptide. This isolated DNA can have a nucleotide sequence that includes, for example, the nucleotide sequence of the *daf*-2 gene shown in Fig. 2B. This isolated DNA can also, in preferred embodiments, complement a *daf-2* mutation in *C. elegans*, an InR mutation in a mouse, or an IGF-1 receptor mutation in a mouse.

The isolated DNA encoding a DAF-2 polypeptide can be included in a vector, such as a vector that is capable of directing the expression of the protein encoded by the DNA in a vector-containing cell. The isolated DNA in the vector can be operatively linked to a promoter, for example, a promoter selected from the group consisting of *daf-2*, *age-1*, *daf-16*, *daf-1*, *daf-4*, *daf-3*, and *akt* promoters. The isolated DNA encoding a DAF-2 polypeptide, or a vector including this DNA, can be contained in a cell, such as a bacterial, mammalian, or nematode cell.

Also included in the invention is a method of producing a recombinant DAF-2 polypeptide, and a DAF-2 polypeptide produced by this method. This method involves (a) providing a cell transformed with isolated DNA that (i) encodes a DAF-2 polypeptide, and (ii) is positioned for expression in the cell, under conditions for expressing the isolated DNA, and (b) isolating the recombinant DAF-2 polypeptide.

A substantially pure antibody, such as a monoclonal or polyclonal antibody, that specifically recognizes and binds a DAF-2 polypeptide is also included in the invention.

The invention also features a method of detecting a gene, or a portion of a gene, that is found in a human cell and has sequence identity to the *daf-2* sequence of Fig. 2B. In this method, isolated DNA encoding a DAF-2 polypeptide, a portion of such DNA greater than about 12 residues in length, or a degenerate oligonucleotide corresponding to SEQ ID NOS: 33, 34, 79, 80, 81, 82, 83, or 84, is contacted with a preparation of DNA from the human cell under hybridization conditions that provide detection of DNA sequences having about 70% or greater nucleic acid sequence identity to the *daf-2* sequence of Fig. 2B. This method can also include a step of testing the gene, or portion thereof, for the ability to functionally complement a *C. elegans daf-2* mutant.

Another method included in the invention is a method of isolating a gene, or a portion of a gene, that is found in a human cell and has at least 90% nucleic acid sequence identity to a sequence encoding SEQ ID NOS: 33, 34, 79, 80, 81, 82, 83, or 84. This method involves (a) amplifying by PCR the human gene, or portion thereof, using oligonucleotide primers that (i) are each greater than about 12 residues in length, and (ii) each have regions of complementarity to opposite DNA strands in a region of the nucleotide sequence of Fig. 2B, and (b) isolating the human gene, or portion thereof. This method can also include a step of testing the gene, or portion thereof, for the ability to functionally complement a *C. elegans daf-2* mutant.

In another aspect, the invention features a substantially pure preparation of a DAF-3 polypeptide, which can be derived from an animal (for example, a mammal, such as a human, or an invertebrate, such as *C. elegans*). In a preferred embodiment, the polypeptide is a SMAD protein. In other preferred embodiments, the polypeptide is capable of binding and interacting with a nematode DAF-1, DAF-4, DAF-8, DAF-14, or DAF-16 polypeptide.

The invention also features isolated DNA encoding a DAF-3

polypeptide. This isolated DNA can have a sequence that includes, for example, the nucleotide sequence of a *daf-3* gene shown in Figs. 11A, 11B, or 11C. This isolated DNA can also, in preferred embodiments, complement a *daf-3* mutation in *C. elegans* or complement a *daf-3* knockout mouse.

The isolated DNA encoding a DAF-3 polypeptide can be included in a vector, such as a vector that is capable of directing the expression of the protein encoded by the DNA in a vector-containing cell. The isolated DNA in the vector can be operatively linked to a promoter, for example, a promoter selected from the group consisting of *daf-3*, *daf-4*, *daf-16*, *daf-2*, *age-1*, and *akt* promoters. The isolated DNA encoding a DAF-3 polypeptide, or a vector including this DNA, can be contained in a cell, such as a bacterial, mammalian, or nematode cell.

Also included in the invention is a method of producing a recombinant DAF-3 polypeptide, and a DAF-3 polypeptide produced by this method. This method involves (a) providing a cell transformed with isolated DNA that (i) encodes a DAF-3 polypeptide, and (ii) is positioned for expression in the cell, under conditions for expressing the isolated DNA, and (b) isolating the recombinant DAF-3 polypeptide.

A substantially pure antibody, such as a monoclonal or polyclonal antibody, that specifically recognizes and binds a DAF-3 polypeptide is also included in the invention.

The invention also features a method of detecting a gene, or a portion of a gene, that is found in a human cell and has sequence identity to any of the *daf-3* sequences of Figs. 11A, 11B, or 11C. In this method, isolated DNA encoding a DAF-3 polypeptide, a portion of such DNA that is greater than about 12 residues in length, or a degenerate oligonucleotide corresponding to SEQ ID NOS: 35, 36, or 85, is contacted with a preparation of DNA from the human cell under hybridization conditions that provide detection of DNA

sequences having about 70% or greater nucleic acid sequence identity to any of the *daf-3* sequences of Figs. 11A, 11B, or 11C. This method can also include a step of testing the gene, or portion thereof, for the ability to functionally complement a *C. elegans daf-3* mutant.

Another method included in the invention is a method of isolating a gene, or a portion thereof, that is found in a human cell and has at least 90% nucleic acid sequence identity to a sequence encoding SEQ ID NOS: 35, 36, or 85. This method includes (a) amplifying by PCR the human gene, or portion thereof, using oligonucleotide primers that (i) are each greater than about 12 residues in length, and (ii) each have regions of complementarity to opposite DNA strands in a region of any of the nucleotide sequences of Figs. 11A, 11B, or 11C, and (b) isolating the human gene, or portion thereof. This method can also include a step of testing the gene, or portion thereof, for the ability to functionally complement a *C. elegans daf-3* mutant.

In yet another aspect, the invention features a substantially pure preparation of DAF-16 polypeptide, which can be derived from an animal (for example, a mammal, such as a human, or an invertebrate, such as *C. elegans*). In a preferred embodiment, the polypeptide is a forkhead transcription factor that binds DNA. In other preferred embodiments, the polypeptide is capable of interacting with a polypeptide selected from the group consisting of DAF-3, DAF-8, and DAF-14.

The invention also features isolated DNA encoding a DAF-16 polypeptide. This isolated DNA can have a sequence that includes, for example, the sequence of the *daf*-16 gene shown in Figs. 13A or 13B. This isolated DNA can also, in preferred embodiments, complement a *daf*-16 mutation in *C. elegans*, or complement an FKHR or AFX mutation in a mouse.

The isolated DNA encoding a DAF-16 polypeptide can be included in a vector, such as a vector that is capable of directing the expression of the protein

encoded by the DNA in a vector-containing cell. The isolated DNA in the vector can be operatively linked to a promoter, for example, a promoter selected from the group consisting of *daf-2*, *age-1*, *daf-16*, *daf-3*, *daf-4*, and *akt* promoters. The isolated DNA encoding a DAF-16 polypeptide, or a vector containing this DNA, can be contained in a cell, such as a bacterial, mammalian, or nematode cell.

Also included in the invention is a method for producing a recombinant DAF-16 polypeptide, and a DAF-16 polypeptide produced by this method. This method involves (a) providing a cell transformed with purified DNA that (i) encodes a DAF-16 polypeptide, and (ii) is positioned for expression in the cell, under conditions for expressing the isolated DNA, and (b) isolating the recombinant DAF-16 polypeptide.

A substantially pure antibody, such as a monoclonal or polyclonal antibody, that specifically recognizes and binds a DAF-16 polypeptide is also included in the invention.

The invention also features a method of detecting a gene, or a portion of a gene, that is found in a human cell and has sequence identity to the *daf-16* sequence of Figs. 13A or 13B. In this method, isolated DNA encoding a DAF-16 polypeptide, a portion of such DNA that is greater than about 12 residues in length, or a degenerate oligonucleotide corresponding to SEQ ID NO: 54, 55, 56, or 57, is contacted with a preparation of DNA from the human cell under hybridization conditions that provide detection of DNA sequences having about 70% or greater nucleic acid sequence identity to the *daf-16* sequence of Figs. 13A or 13B. This method can also include a step of testing the gene, or portion of the gene, for the ability to functionally complement a *C. elegans daf-16* mutant.

Another method included in the invention is a method of isolating a gene, or a portion of a gene, that is found in a human cell and has at least 90%

nucleic acid sequence identity to a sequence encoding SEQ ID NO: 54, 55, 56, or 57. This method involves (a) amplifying by PCR the human gene, or portion thereof, using oligonucleotide primers that (i) are each greater than about 12 residues in length, and (ii) each have regions of complementarity to opposite DNA strands in a region of the nucleotide sequence of Figs. 13A or 13B, and (b) isolating the human gene, or portion thereof. This method can also include a step of testing the gene, or portion thereof, for the ability to functionally complement a *C. elegans daf-16* mutant.

In another aspect, the invention features a method of determining whether a human gene is involved in an impaired glucose tolerance condition (for example, a condition involving atherosclerosis) or obesity. This method involves (a) providing a nematode having a mutation in a *daf* or *age* gene, and (b) expressing in the nematode the human gene, which is operatively linked to a nematode gene promoter. Complementation of the *daf* or *age* mutation in the nematode is indicative of a human gene that is involved in an impaired glucose tolerance condition or obesity. In preferred embodiments, the nematode gene promoter is selected from the group consisting of *daf-1*, *daf-3*, *daf-4*, *daf-2*, *age-1*, and *akt* gene promoters. In other preferred embodiments, the *daf* mutation is selected from the group consisting of *daf-2*, *daf-3*, *daf-1*, *daf-4*, *daf-7*, *daf-8*, *daf-11*, *daf-12*, *daf-14*, and *daf-16* mutations. In yet another preferred embodiment, the mutation can also be found in the *age-1* gene.

In further aspects, the invention features methods for diagnosing an impaired glucose tolerance condition (for example, Type II diabetes or a condition involving atherosclerosis), or a propensity for such a condition, in a patient. One such method includes analyzing the DNA of the patient to determine whether the DNA contains a mutation in a *daf* gene. Identification of such a mutation indicates that the patient has an impaired glucose tolerance condition or a propensity for such a condition. The analysis in this method can

be carried out, for example, by nucleotide sequencing or RFLP analysis. The analysis can also include amplifying (for example, by PCR or reverse transcriptase PCR) the gene (for example, a human gene), or a fragment thereof, using primers, and analyzing the amplified gene, or a fragment thereof, for the presence of the mutation. In preferred embodiments, the *daf* gene analyzed in this method is, for example, a *daf-1*, *daf-2*, *daf-3* *daf-4*, *daf-7*, *daf-8*, *daf-11*, *daf-12*, *daf-14*, or *daf-16* coding sequence, or the d*af* gene is FKHR or AFX.

Another method for diagnosing an impaired glucose tolerance condition, such as Type II diabetes, or a propensity for such a condition, in a patient, includes analyzing the DNA of the patient to determine whether the DNA contains a mutation in an *age* gene. Identification of such a mutation indicates that the patient has an impaired glucose tolerance condition or a propensity for such a condition. The analysis in this method can be carried out, for example, by nucleotide sequencing or RFLP analysis. The analysis can also include amplifying (for example, by PCR or reverse transcriptase PCR) the gene (for example, a human gene), or a fragment thereof, using primers and analyzing the amplified gene, or fragment thereof, for the presence of the mutation. In a preferred embodiment, the *age* gene is an *age-1* coding sequence.

Yet another method for diagnosing an impaired glucose tolerance condition, such as Type II diabetes or a condition that involves atherosclerosis, or a propensity for such a condition, in a patient, includes analyzing the DNA of the patient to determine whether the DNA contains a mutation in an *akt* gene. Identification of such a mutation indicates that the patient has an impaired glucose tolerance condition (for example, Type II diabetes) or a propensity for such a condition (for example, a pre-diabetic condition). The analysis in this method can be carried out, for example, by nucleotide sequencing or RFLP analysis. The analysis can also include amplifying (for

example, by PCR or reverse transcriptase PCR) the gene (for example, a human gene), or a fragment thereof, using primers and analyzing the amplified gene, or fragment thereof, for the presence of the mutation.

The invention also includes kits for use in the diagnosis of an impaired glucose tolerance condition, or a propensity for such a condition, in a patient. One such kit includes a PCR primer complementary to a *daf* nucleic acid sequence and instructions for diagnosing an impaired glucose tolerance condition or a propensity for such a condition. Another kit includes a PCR primer complementary to an *age* nucleic acid sequence and instructions for diagnosing an impaired glucose tolerance condition or a propensity for such a condition. Yet another kit includes a PCR primer complementary to an *akt* nucleic acid sequence and instructions for diagnosing an impaired glucose tolerance condition or a propensity for such a condition.

In another aspect, the invention features methods for ameliorating or delaying the onset of an impaired glucose tolerance condition (for example, Type II diabetes) in a patient. In one such method a therapeutically effective amount of a DAF polypeptide (for example, the human or nematode DAF-7 polypeptide) is administered to the patient. In another method, which can be used, for example, in the case of a condition involving atherosclerosis, a therapeutically effective amount of a compound that is capable of inhibiting the activity of a DAF-16 or DAF-3 polypeptide is administered to the patient. In yet another method, a therapeutically effective amount of a compound that activates a DAF-1, DAF-4, DAF-8, DAF-11, or DAF-14 polypeptide is administered to the patient.

Another aspect of the invention provides methods for ameliorating or preventing obesity (for example, obesity associated with Type II diabetes) in a patient. One such method involves administering to the patient a therapeutically effective amount of a DAF polypeptide, such as a human or

nematode DAF-7 polypeptide. Another such method involves administering to the patient a therapeutically effective amount of a compound that is capable of inhibiting the activity of a DAF-16 or DAF-3 polypeptide.

Yet another aspect of the invention features a transgenic, non-human animal, such as a mouse or a nematode, whose germ cells and somatic cells contain a transgene coding for a mutant DAF polypeptide, for example, a mutant DAF polypeptide that is derived from a human. In preferred embodiments, the mutant DAF polypeptide is a DAF-1, DAF-2, DAF-3, DAF-4, DAF-7, DAF-8, DAF-11, DAF-12, DAF-14, or DAF-16 polypeptide. In another preferred embodiment, the transgene includes a knockout mutation.

In a related aspect, the invention features a transgenic, non-human animal, such as a mouse or a nematode, whose germ cells and somatic cells contain a transgene coding for a mutant AGE polypeptide, for example, a mutant AGE polypeptide derived from a human. In a preferred embodiment, the mutant AGE polypeptide is an AGE-1 polypeptide. In another preferred embodiment, the transgene includes a knockout mutation.

In yet another aspect, the invention features a transgenic, non-human animal, such as a mouse or a nematode, whose germ cells and somatic cells contain a transgene coding for a mutant AKT polypeptide, for example, a mutant AKT polypeptide derived from a human. In a preferred embodiment, the transgene includes a knockout mutation.

In related aspects, the invention features cells (for example, cells isolated from a mammal, such as mouse, human, or nematode cells) isolated from the transgenic animals described above.

The invention also includes methods for producing transgenic, non-human animals. For example, the invention includes a method for producing a transgenic, non-human animal that lacks an endogenous *daf* gene and is capable of expressing a human DAF polypeptide. This method involves (a) providing a

transgenic, non-human animal whose germ cells and somatic cells contain a mutation in a *daf* gene, and (b) introducing a transgene that (i) encodes a human DAF polypeptide, and (ii) is capable of expressing the human polypeptide, into an embryonal cell of the non-human animal.

Another method included in the invention can be used for producing a transgenic, non-human animal that lacks an endogenous *age* gene and is capable of expressing a human AGE polypeptide. This method involves (a) providing a transgenic, non-human animal whose germ cells and somatic cells contain a mutation in an *age* gene, and (b) introducing a transgene that (i) encodes a human AGE polypeptide, and (ii) is capable of expressing the human polypeptide, into an embryonal cell of the non-human animal.

Similarly, the invention includes a method for producing a transgenic, non-human animal that lacks an endogenous *akt* gene and is capable of expressing of expressing a human AKT polypeptide. This method involves (a) providing a transgenic, non-human animal whose germ cells and somatic cells contain a mutation in an *akt* gene, and (b) introducing a transgene that (i) encodes a human AKT polypeptide, and (ii) is capable of expressing the human polypeptide, into an embryonal cell of the non-human animal.

Another aspect of the invention features a method of screening for a compound that increases the activity of a DAF polypeptide. This method includes (a) exposing a non-human transgenic animal whose germ cells and somatic cells contain a transgene coding for a mutant DAF polypeptide to a candidate compound, and (b) determining the activity of the DAF polypeptide in the transgenic animal. An increase in DAF polypeptide activity, as compared to untreated controls, is indicative of a compound that is capable of increasing DAF polypeptide activity. In preferred embodiments, the compound can be used to treat an impaired glucose tolerance condition or obesity.

In a related aspect, the invention features a method of screening for a

compound that decreases the activity of a DAF polypeptide. This method includes (a) exposing a non-human transgenic animal whose germ cells and somatic cells contain a transgene coding for a mutant DAF polypeptide to a candidate compound, and (b) determining the activity of the DAF polypeptide in the transgenic animal. A decrease in DAF polypeptide activity, as compared to untreated controls, is indicative of a compound that is capable of decreasing DAF polypeptide activity. In preferred embodiments, the compound can be used to treat an impaired glucose tolerance condition, obesity, or atherosclerosis. In other preferred embodiments, the compound decreases the activity of DAF-3 or DAF-16.

In another aspect, the invention features a method of screening for a compound that increases the activity of an AGE polypeptide. This method includes (a) exposing a non-human transgenic animal whose germ cells and somatic cells contain a transgene coding for a mutant AGE polypeptide to a candidate compound, and (b) determining the activity of the AGE polypeptide in the transgenic animal. An increase in AGE polypeptide activity, as compared to untreated controls, is indicative of a compound that is capable of increasing AGE polypeptide activity. In preferred embodiments, the compound can be used to treat an impaired glucose tolerance condition, obesity, or atherosclerosis.

In a related aspect, the invention features a method of screening for a compound that decreases the activity of a AGE polypeptide. This method includes (a) exposing a non-human, transgenic animal whose germ cells and somatic cells contain a transgene coding for a mutant AGE polypeptide to a candidate compound, and (b) determining the activity of the AGE polypeptide in the transgenic animal. A decrease in AGE polypeptide activity, as compared to untreated controls, is indicative of a compound that is capable of decreasing AGE polypeptide activity. In preferred embodiments, the compound can be

used to treat an impaired glucose tolerance condition, obesity, or atherosclerosis. In another preferred embodiment, the AGE polypeptide is AGE-1.

In another aspect, the invention features a method of screening for a compound that increases the activity of an AKT polypeptide. This method includes (a) exposing a transgenic, non-human animal whose germ cells and somatic cells contain a transgene coding for a mutant AKT polypeptide to a candidate compound, and (b) determining the activity of the AKT polypeptide in the transgenic animal. An increase in AKT polypeptide activity, as compared to untreated controls, is indicative of a compound that is capable of increasing AKT polypeptide activity. In preferred embodiments, the compound can be used to treat an impaired glucose tolerance condition, obesity, or atherosclerosis.

In a related aspect, the invention features a method of screening for a compound that decreases the activity of a AKT polypeptide. This method includes (a) exposing a transgenic, non-human animal whose germ cells and somatic cells contain a transgene coding for a mutant AKT polypeptide to a candidate compound, and (b) determining the activity of the AKT polypeptide in the transgenic animal. A decrease in AKT polypeptide activity, as compared to untreated controls, is indicative of a compound that is capable of decreasing AKT polypeptide activity. In preferred embodiments, the compound can be used to treat an impaired glucose tolerance condition or obesity.

Also included in the invention is a method of screening for a compound that is capable of ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) exposing a transgenic, non-human animal whose germ cells and somatic cells contain a transgene coding for a mutant DAF, AGE, or AKT polypeptide to a candidate compound, and (b) monitoring the blood glucose level of the animal. A compound that promotes maintenance

of a physiologically acceptable level of blood glucose in the animal, as compared to untreated controls, is indicative of a compound that is capable of ameliorating or delaying an impaired glucose tolerance condition. In a preferred embodiment, the compound can be used to treat Type II diabetes.

Another method of screening for a compound that is capable of ameliorating or delaying obesity is also included in the invention. This method involves (a) exposing a transgenic, non-human animal whose germ cells and somatic cells contain a transgene coding for a mutant DAF, AGE, or AKT polypeptide to a candidate compound, and (b) monitoring the adipose tissue of the animal. A compound that promotes maintenance of a physiologically acceptable level of adipose tissue in the animal, as compared to untreated controls, is indicative of a compound that is capable of ameliorating or delaying obesity.

A related method of the invention can be used for screening for a compound that is capable of ameliorating or delaying atherosclerosis. This method involves (a) exposing a transgenic, non-human animal whose germ cells and somatic cells contain a transgene coding for a mutant DAF, AGE, or AKT polypeptide to a candidate compound, and (b) monitoring the adipose tissue of the animal. A compound that promotes maintenance of a physiologically acceptable level of adipose tissue in the animal, as compared to untreated controls, is indicative of a compound that is capable of ameliorating or delaying atherosclerosis.

In another aspect, the invention includes a method for identifying a modulatory compound that is capable of decreasing the expression of a *daf* gene. This method involves (a) providing a cell expressing the *daf* gene, and (b) contacting the cell with a candidate compound. A decrease in *daf* expression following contact with the candidate compound identifies a modulatory compound. In preferred embodiments, the compound can be used

to treat an impaired glucose tolerance condition or obesity. In other preferred embodiments, the compound is capable of decreasing the expression of DAF-3 or DAF-16. This method can be carried out in an animal, such as a nematode.

In a related aspect, the invention includes a method for the identification of a modulatory compound that is capable of increasing the expression of a *daf* gene. This method involves (a) providing a cell expressing the *daf* gene, and (b) contacting the cell with a candidate compound. An increase in *daf* expression following contact with the candidate compound identifies a modulatory compound. In preferred embodiments, the compound can be used to treat an impaired glucose tolerance condition or obesity. In other preferred embodiments, the compound is capable of increasing expression of DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, or DAF-14. This method can be carried out in an animal, such as a nematode.

In another aspect, the invention includes a method for the identification of a modulatory compound that is capable of increasing the expression of an *age-1* gene. This method involves (a) providing a cell expressing the *age-1* gene, and (b) contacting the cell with a candidate compound. An increase in *age-1* expression following contact with the candidate compound identifies a modulatory compound. In preferred embodiments, the compound is capable of treating an impaired glucose tolerance condition or obesity. This method can be carried out in an animal, such as a nematode.

In another aspect, the invention provides a method for identification of a compound that is capable of ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) providing a dauer larvae including a mutation in a *daf* gene, and (b) contacting the dauer larvae with a compound. Release from the dauer larval state is an indication that the compound is capable of ameliorating or delaying an impaired glucose tolerance condition. In a preferred embodiment, the dauer larvae carries a *daf-2*

mutation. In another preferred embodiment, the dauer larvae is from *C. elegans*. In yet another embodiment, the impaired glucose tolerance condition involves obesity or atherosclerosis.

In a related aspect, the invention provides a method for identification of a compound that is capable of ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) providing a dauer larvae including a mutation in an *age-1* gene, and (b) contacting the dauer larvae with a compound. Release from the dauer larval state is an indication that the compound is capable of ameliorating or delaying an impaired glucose tolerance condition. In a preferred embodiment, the dauer larvae carries an *age-1* mutation. In another preferred embodiment, the dauer larvae is from *C. elegans*. In yet another preferred embodiment, the impaired glucose tolerance condition involves obesity or atherosclerosis.

In another related aspect, the invention provides a method for the identification of a compound that is capable of ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) providing a dauer larvae including a mutation in an *akt* gene, and (b) contacting the dauer larvae with a compound. Release from the dauer larval state is an indication that the compound is capable of ameliorating or delaying an impaired glucose tolerance condition. In a preferred embodiment, the dauer larvae is from *C. elegans*. In another preferred embodiment, the impaired glucose tolerance condition involves obesity or atherosclerosis.

In another aspect, the invention provides a method for the identification of a compound for ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) combining PIP3 and an AKT polypeptide in the presence and absence of a compound under conditions that allow PIP3:AKT complex formation, (b) identifying a compound that is capable of decreasing the formation of the PIP3:AKT complex, and (c) determining

whether the compound identified in step (b) is capable of increasing AKT activity. An increase in AKT kinase activity is taken as an indication of a compound useful for ameliorating or delaying an impaired glucose tolerance condition.

In yet another aspect, the invention provides a method for the identification of a compound for ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) providing a *daf-7*, *daf-3* mutant nematode, (b) expressing in the cells of the nematode a mammalian DAF-3 polypeptide, whereby the nematode forms a dauer larva, and (c) contacting the dauer larva with a compound. A release from the dauer larval state is an indication that the compound is capable of ameliorating or delaying the glucose intolerance condition.

In a further aspect, the invention features a method for the identification of a compound for ameliorating or delaying an impaired glucose tolerance condition. This method involves (a) providing a *daf-2*, *daf-16* mutant nematode, (b) expressing in the cells of the nematode a mammalian DAF-16 polypeptide, whereby the nematode forms a dauer larva, and (c) contacting the dauer larva with a compound. A release from the dauer larval state is an indication that the compound is capable of ameliorating or delaying the glucose intolerance condition.

In a final aspect, the invention features insulin-like molecules and their use as diagnostic and therapeutic reagents.

As used herein, by a "DAF" polypeptide is meant a polypeptide that functionally complements a *C. elegans daf* mutation and/or that has at least 60%, preferably 75%, and more preferably 90% amino acid sequence identity to a 100 amino acid region (and preferably a conserved domain) of a *C. elegans* DAF polypeptide. Complementation may be assayed in an organism (for example, in a nematode) or in a cell culture system. Complementation may be

partial or complete, but must provide a detectable increase in function (as described herein). DAF polypeptides are encoded by "DAF" genes or nucleic acid sequences.

By an "AGE" polypeptide is meant a polypeptide that functionally complements a *C. elegans age* mutation and/or that has at least 60%, preferably 75%, and more preferably 90% amino acid sequence identity to a 100 amino acid region (and preferably a conserved domain) of a *C. elegans* AGE polypeptide. Complementation may be assayed in an organism (for example, in a nematode) or in a cell culture system. Complementation may be partial or complete, but must provide a detectable increase in a known AGE function. AGE polypeptides are encoded by "AGE" genes or nucleic acid sequences.

As used herein, by an "AKT" polypeptide is meant a polypeptide that functionally complements a *C. elegans akt* mutation and/or that possess at least 64% amino acid sequence identity to SEQ ID NO: 60, at least 71% amino acid sequence identity to SEQ ID NO: 61, at least 79% amino acid sequence identity to SEQ ID NO: 62, at least 63% amino acid sequence identity to SEQ ID NO: 63, at least 48% amino acid sequence identity to SEQ ID NO: 64, at least 70% amino acid sequence identity to SEQ ID NO: 65, at least 64% amino acid sequence identity to SEQ ID NO: 66, at least 67% amino acid sequence identity to SEQ ID NO: 67, or a combination thereof. Complementation may be assayed in an organism (for example, in a nematode) or in a cell culture system. Complementation may be partial or complete, but must provide a detectable increase in a known AKT function. AKT polypeptides are encoded by "AKT" genes or nucleic acid sequences.

By a "DAF-2 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-2* mutation and/or that possesses at least 61% amino acid sequence identity to SEQ ID NO: 33, at least 31% amino acid sequence identity to SEQ ID NO: 34, at least 43% amino acid sequence identity

to SEQ ID NO: 79, at least 35% amino acid sequence identity to SEQ ID NO: 80, at least 35% amino acid sequence identity to SEQ ID NO: 81, at least 48% amino acid sequence identity to SEQ ID NO: 82, at least 43% amino acid sequence identity to SEQ ID NO: 83, at least 40% amino acid sequence identity to SEQ ID NO: 84, or a combination thereof. Preferably, a DAF-2 polypeptide includes an aspartic acid, a proline, a proline, a serine, an alanine, an aspartic acid, a cysteine, or a proline at amino acid positions corresponding to *C. elegans* DAF-2 amino acids 1252, 1312, 1343, 347, 451, 458, 526, 279, and 348 respectively, or a combination thereof.

By a "DAF-3 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-3* mutation and/or that possesses at least 60% amino acid sequence identity to SEQ ID NO: 35, at least 38% amino acid sequence identity to SEQ ID NO: 36, at least 47% amino acid sequence identity to SEQ ID NO: 85, or a combination thereof. Preferably, a DAF-3 polypeptide includes a proline or a glycine at amino acid positions corresponding to *C. elegans daf-3* amino acids at positions 200 (proline) and/or 620 (glycine) in Fig. 12A, respectively, or a combination thereof. For example, the polypeptide may include a proline in the motif GRKGFPHV or a glycine in the motif RXXIXXG (where X is any amino acid).

By a "DAF-16 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-16* mutation and/or that possesses at least 71% amino acid sequence identity to SEQ ID NO: 54, at least 35% amino acid sequence identity to SEQ ID NO: 55, at least 65% amino acid sequence identity to SEQ ID NO: 56, at least 53% amino acid sequence identity to SEQ ID NO: 57, or a combination thereof. In addition, a DAF-16 polypeptide preferably includes a serine residue in the conserved motif WKNSIRH (SEQ ID NO: 59).

By a "DAF-7 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-7* mutation and/or that possesses at least 29%

amino acid sequence identity to SEQ ID NO: 26, at least 66% amino acid sequence identity to SEQ ID NO: 27, at least 45% amino acid sequence identity to SEQ ID NO: 28, at least 33% amino acid sequence identity to SEQ ID NO: 29, at least 56% amino acid sequence identity to SEQ ID NO: 30, at least 75% sequence identity to SEQ ID No: 51, or a combination thereof. Preferably, a DAF-7 polypeptide includes a proline or a glycine at amino acid positions corresponding to *C. elegans daf-7* amino acids 271 and 280, respectively, or a combination thereof.

By a "DAF-8 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-8* mutation and/or that possesses at least 46% amino acid sequence identity to SEQ ID NO: 23, at least 45% amino acid sequence identity to SEQ ID NO: 24, at least 36% amino acid sequence identity to SEQ ID NO: 25, or a combination thereof.

By an "AGE-1 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans age-1* mutation (previously known as a *daf-23* mutation) and/or that possesses at least 40% amino acid sequence identity to SEQ ID NO: 17, at least 45% amino acid sequence identity to SEQ ID NO: 18, at least 30% amino acid sequence identity to SEQ ID NO: 19, at least 24% amino acid sequence identity to SEQ ID NO: 38, or a combination thereof. Preferably, an AGE-1 polypeptide includes an alanine at amino acid positions corresponding to *C. elegans age-1* amino acids 845.

By a "DAF-1 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-1* mutation and/or that possesses at least 45% amino acid sequence identity to SEQ ID NO: 13, at least 35% amino acid sequence identity to SEQ ID NO: 14, at least 65% amino acid sequence identity to SEQ ID NO: 15, at least 25% amino acid sequence identity to SEQ ID NO: 16, or a combination thereof. Preferably, a DAF-1 polypeptide includes a proline at the amino acid position corresponding to *C. elegans* DAF-1 amino

acid 546.

By a "DAF-4 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-4* mutation and/or that possesses at least 45% amino acid sequence identity to SEQ ID NO: 20, at least 40% amino acid sequence identity to SEQ ID NO: 21, at least 44% amino acid sequence identity to SEQ ID NO: 22, or a combination thereof.

By a "DAF-11 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-11* mutation and/or that possesses at least 40% amino acid sequence identity to SEQ ID NO: 75, at least 43% amino acid sequence identity to SEQ ID NO: 76, at least 36% amino acid sequence identity to SEQ ID NO: 77, at least 65% amino acid sequence identity to SEQ ID NO: 78, or a combination thereof.

By a "DAF-12 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-12* mutation and/or that possesses at least 42% amino acid sequence identity to SEQ ID NO: 72, at least 58% amino acid sequence identity to SEQ ID NO: 73, at least 34% amino acid sequence identity to SEQ ID NO: 74, or a combination thereof.

By a "DAF-14 polypeptide" is meant a polypeptide that complements (as defined above) a *C. elegans daf-14* mutation and/or that possesses at least 48% amino acid sequence identity to SEQ ID NO: 68, at least 37% amino acid sequence identity to SEQ ID NO: 69, at least 48% amino acid sequence identity to SEQ ID NO: 70, at least 37% amino acid sequence identity to SEQ ID NO: 71, or a combination thereof.

By "insulin receptor activity" is meant any activity exhibited by an insulin receptor and measured by either (i) activation of insulin receptor substrate-1 (IRS-1) phosphorylation and recruitment of PI-3 kinase, (ii) activation of glucose transporter (Glut 4) fusion with a cellular membrane and concomitant glucose uptake, or (iii) activation of glycogen and/or fat synthesis

and concomitant inhibition of gluconcogenesis or lipolysis or both.

By "insulin receptor related activity" is meant any activity not directly attributable to the insulin receptor but that is measured by an activation of IRS-1 phosphorylation and recruitment of PI3-kinase.

By "IGF-1 receptor activity" is meant any activity exhibited by an insulin-like growth factor-1 receptor and measured by (i) activation of IRS-1 phosphorylation and recruitment of PI-3 kinase, (ii) activation of cell division in NIH3T3 cells (e.g., as described in Gronborg et al., J. Biol. Chem. 268: 23435-23440, 1993), or (iii) activation of bone growth in, for example, the mouse model.

By "SMAD protein" is meant a protein that is capable of coupling to TGF-β type ser/thr receptors. Smad proteins typically contain a smad conserved motif as described by Derynk et al. (*Cell* 87: 173, 1996). Exemplary smad proteins include, without limitation, DAF-3, MADR-2, MAD, DPC-4, and Sma-2.

By "AKT activity" is meant any activity exhibited by an AKT polypeptide and measured by phosphatidylinositol-regulated increases in serine phosphorylation of GSK-3 or activation of non-dauer growth in *C. elegans akt* mutants.

By "impaired glucose tolerance condition" is meant any condition in which blood sugar levels are inappropriately elevated or lack normal metabolic regulation. Examples of such conditions include, without limitation, Type I diabetes, Type II diabetes, and gestational diabetes, and may be associated with obesity and atherosclerosis.

By "protein" or "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation or phosphorylation).

By "substantially pure" is meant a preparation which is at least 60% by

weight (dry weight) the compound of interest, e.g., any of the polypeptides of the invention such as the DAF-2, DAF-3, or DAF-16 polypeptides or DAF-2, DAF-3, or DAF-16-specific antibodies. Preferably the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate method, e.g., column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

By "isolated DNA" is meant DNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally-occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By a "substantially identical" polypeptide sequence is meant an amino acid sequence which differs only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for glycine, arginine for lysine, etc.) or by one or more non-conservative substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the function of the polypeptide (assayed, e.g., as described herein).

Preferably, such a sequence is at least 75%, more preferably 85%, and most preferably 95% identical at the amino acid level to the sequence used for comparison.

Homology is typically measured using sequence analysis software (e.g.,

Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705 or BLAST software available from the National Library of Medicine). Examples of useful software include the programs, Pileup and PrettyBox. Such software matches similar sequences by assigning degrees of homology to various substitutions, deletions, substitutions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

By a "substantially identical" nucleic acid is meant a nucleic acid sequence which encodes a polypeptide differing only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for glycine, arginine for lysine, etc.) or by one or more non-conservative substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the function of the polypeptide (assayed, e.g., as described herein). Preferably, the encoded sequence is at least 75%, more preferably 85%, and most preferably 95% identical at the amino acid level to the sequence of comparison. If nucleic acid sequences are compared a "substantially identical" nucleic acid sequence is one which is at least 85%, more preferably 90%, and most preferably 95% identical to the sequence of comparison. The length of nucleic acid sequence comparison will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 110 nucleotides. Again, homology is typically measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705).

By "positioned for expression" is meant that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of any of the polypeptides disclosed herein including, but not limited to, DAF-2, DAF-3, and DAF-16 and any human homolog thereof).

By "purified antibody" is meant antibody which is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, antibody.

By "specifically binds" is meant an antibody which recognizes and binds a polypeptide of the invention (e.g., DAF-2, DAF-3, and DAF-16) but which does not substantially recognize and bind other molecules in a sample (e.g., a biological sample) which naturally includes a polypeptide of the invention. An antibody which "specifically binds" such a polypeptide is sufficient to detect protein product in such a biological sample using one or more of the standard immunological techniques available to those in the art (for example, Western blotting or immunoprecipitation).

By "immunological methods" is meant any assay involving antibody-based detection techniques including, without limitation, Western blotting, immunoprecipitation, and direct and competitive ELISA and RIA techniques.

By "means for detecting" is meant any one or a series of components that sufficiently indicate a detection event of interest. Such means involve at least one label that may be assayed or observed, including, without limitation, radioactive, fluorescent, and chemiluminescent labels.

By "hybridization techniques" is meant any detection assay involving specific interactions (based on complementarity) between nucleic acid strands, including DNA-DNA, RNA-RNA, and DNA-RNA interactions. Such

hybridization techniques may, if desired, include a PCR amplification step.

By a "modulatory compound", as used herein, is meant any compound capable of either decreasing DAF-3 and DAF-16 expression (i.e., at the level of transcription, translation, or post-translation) or decreasing DAF-3 and DAF-16 protein levels or activity. Also included are compounds capable of either increasing DAF-1, DAF-2, DAF-4, DAF-8, DAF-7, DAF-11, DAF-14, AGE-1, and AKT expression (i.e., at the level of transcription, translation, or post-translation) or increasing DAF-1, DAF-2, DAF-4, DAF-8, DAF-7, DAF-11, DAF-14, AGE-1, and AKT protein levels or their corresponding activities.

By "complementation" is meant an improvement of a genetic defect or mutation. In one example, complementation of a genetic defect in a *daf*, *age*, or *akt* gene can be carried out by providing the wild-type *daf*, *age*, or *akt* genes, respectively. Complementation is generally accomplished by expressing the wild-type version of the protein in a host cell or animal bearing a mutant or inactive version of the gene.

Other features and advantages of the invention will be apparent from the following detailed description thereof, and from the claims.

Detailed Description

The drawings will first be described.

Drawings

Fig. 1 shows the genetic and physical map of *C. elegans daf-2*. The top panel shows the genetic map of *daf-2*. *daf-2* maps on the left arm of chromosome III 11.4 map units to the right of *dpy-1* and 1.6 map units to the left of *ben-1* (ACeDB). The middle panel shows the physical map of *daf-2*. *daf-2* maps between *mg*P34 and *mg*P44 in a region not covered by cosmid clones but covered by YAC Y53G8. Cosmids from the approximate *daf-2* genetic location detect RFLPs between *C. elegans* strains Bristol N2 and

Bergerac RC301. *mgP31* on cosmid T21A6 is a HindIII RFLP: 5.3 kb in Bristol, 4.5 kb in RC301. *mgP33* on cosmid T02B2 is a HindIII RFLP: 9 kb in Bristol, 8 kb in RC301. *mgP34* on cosmid R10F2 is an EcoRI RFLP: 4.1 and 2.8 kb in Bristol, 3.6 kb in RC301. *mgP44* on cosmid R07G11 is a complex EcoRI RFLP: 2.9 kb, 2.4 kb, 1.9 kb and 1.7kb in Bristol; 3.6kb, 2.5kb and 1.6kb in RC301. *mgP35* on cosmid T10D5 is a StyI RFLP: 5.4 kb in Bristol, 5.8 kb in RC301. *mgP32* on cosmid C42B8 is a StyI RFLP: 2.8 kb in Bristol; 2.9kb in RC301. *mgP48* detected with *daf-2* probe (nt 1277-2126 and 3747-4650) is a HindIII RFLP: 4.3kb and 7kb in Bristol and 4.1kb and 6.2kb in RC301. Thirty-one out of thirty-three Dpy-non-Daf recombinants carry the RC301 allele of *mgP34* whereas all thirty-three recombinants in this interval carry the RC301 allele of *mgP44*, mapping *daf-2* 0.69 map units to the right of *mgP34* and to the left of *mgP44*. Fourteen out of twenty-four Ben-non-Daf recombinants carry the RC301 *mgP44* allele whereas all of these recombinants carry the RC301 allele of *mgP34*, mapping *daf-2* 0.66 map units to the left of *mgP44*.

Y53G8 YAC DNA was isolated from CHEF gels as described in Ausubel et al. (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y., 1990), labeled , and shown to hybridize to multiple restriction fragments from cosmids bearing *mgP34* and *mgP44*. A probe from the insulin receptor homolog on Y53G8 detects the *mgP48* RFLP between N2 and RC301. All thirty-three Dpy-non-Daf and all twenty-four Ben-non-Daf recombinants described above carry the RC301 allele of *mgP48*, indicating that *daf-2* could not be separated from this insulin receptor gene by these fifty-seven recombination events in a thirteen map unit interval.

The bottom panel shows the structure of *daf-2* cDNA. The *daf-2* cDNA was amplified from a cDNA library constructed according to standard methods by PCR using internal primers derived from the genomic shotgun sequences,

vector sequence primers (for 3' end) and an SL1 transspliced leader PCR primer (M. Krause, In: *Methods Cell Biol.*, vol. 48, pp. 483-512, H. F. Epstein and D. C. Shakes, eds., Academic Press, San Diego, CA, 1995). To isolate a cDNA, pooled plasmid DNA from 106 clones of a 107 clone complexity cDNA library was used as a PCR template. To obtain a *daf-2* cDNA 3' end, *daf-2* internal primer CGCTACGGCAAAAAAGTGAA (SEQ ID NO: 1) in the kinase domain and a cloning vector primer CGATGATGAAGATACCCC (SEQ ID NO: 2) were used in a nested PCR reaction with adjacent internal primers. For the cDNA fragment from the ligand-binding domain to the kinase domain, PCR was carried out with TGATGCGAACGGCGATCGAT (SEQ ID NO: 3) and ACGCTGGATCATCTACATTA (SEQ ID NO: 4) primers. For the *daf*-2 5' end, SL1 primer GGTTTAATTACCCAAGTTTGAG (SEQ ID NO: 5) and one internal *daf-2* primer GCTCACGGGTCACACAACGA (SEQ ID NO: 6) were used in a nested PCR reaction with adjacent internal primers. Using PCR to amplify genomic DNA from a set of 20 *daf-2* mutants, we searched for *daf-2* mutations in a 0.8 kb region of the ligand binding domain and in a 0.9 kb region of the kinase domain. For sequencing the ligand-binding domain PCR primers TGATGCGAACGGCGATCGAT (SEQ ID NO: 7) and TGAGGGCCAACTAAAGAAGAC (SEQ ID NO: 8) were used. In the kinase domain primers CGCTACGGCAAAAAAGTGAA (SEQ ID NO: 9) and GACGATCCCGAGGTGAGTAT (SEQ ID NO: 10) were used. The presence of an SL1 spliced leader sequence indicates a full length *daf-2* cDNA. The predicted ORF is shown as a box; 5' and 3' UTRs are shown as thick bars. The predicted DAF-2 initiator methionine at base 486 is preceded by an in frame stop codon 63 bases upstream. The predicted DAF-2 stop codon is found at base 5658. No consensus polyadenylation signal was found in the cDNA nor in genomic shotgun sequence #00678, which extends 302 bp further downstream. The initial insulin receptor homolog shotgun sequences are shown as thin bars

above the box.

Introns were detected by a combination of *in silico* genomic and cDNA sequence comparison, and by comparison of PCR products derived from cDNA and genomic DNA templates. The open triangles over a vertical bar indicate positions of the detected exon/intron boundaries. All the intron donor sites have GT consensus and the acceptor sites have AG consensus (Krause, 1995 *supra*). The triangles without a vertical bar indicate the approximate intron locations determined by comparison of PCR products using genomic DNA or cDNA as a template. Intron lengths were estimated by comparison of the PCR product size using cDNA or genomic DNA templates. Genomic regions corresponding to some of the introns could not be PCR amplified suggesting that these introns are long. The minimum *daf-2* gene size based on this analysis is 33 kb.

Fig. 2A shows the predicted *C. elegans* DAF-2 amino acid sequence. The predicted cysteine-rich region (amino acids 207-372) and tyrosine kinase domain (amino acids 1124-1398) are boxed. The signal peptide (amino acids 1-20), proteolysis site (amino acids 806-809), transmembrane domain (amino acids 1062-1085), and PTB binding motif in the juxtamembrane region (NPEY, amino acids 1103-1106) are underlined. Three DAF-2 tyrosine residues, Y1293, Y1296 and Y1297, in the region corresponding to the insulin receptor kinase Y1158 to Y1163 activation loop are likely to be autophosphorylated, based on the predicted similarity between the DAF-2 and insulin receptor phosphorylation targets (Fig. 2B). Another likely target for DAF-2 autophosphorylation is the Y1106 NPEY motif located in the region corresponding to the insulin receptor juxtamembrane region NPEY motif (at Y972), that has been shown to mediate IRS-1 binding via its PTB domain to the insulin receptor (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). While DAF-2 bears one YXXM motif implicated in coupling to PI 3-kinase,

mammalian IRS-1 and *Drosophila* insulin receptor (Fernandez et al., EMBO J. 14: 3373-3384, 1995) bear multiple YXXM motifs. Although no p85-like adaptor subunit has yet been detected in the *C. elegans* database, the AGE-1 homology to mammalian p110 suggests the existence of a homologous or analogous adaptor (Morris et al., *Nature* 382: 536-539, 1996). In the DAF-2 C-terminal domain, two other tyrosine residues may be autophosphorylated and bound to particular SH2-containing proteins: Y1678 binding to a PLC-g or SHP-2 homolog, and Y1686, perhaps binding to SEM-5 (Fig. 2A) (Songyang et al., *Cell* 72: 767-778, 1993). While mutations in, for example, ras and MAP kinase have not been identified in screens for dauer constitutive or dauer defective mutations, these general signaling pathway proteins may couple to DAF-2 as they couple to insulin signaling in vertebrates (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). The predicted phosphotyrosine residues in juxtamembrane region and the kinase domain activation loop are circled. In the extended C-terminal region, predicted phosphotyrosine residues are also circled and SH2-binding sites are underlined (see below).

Fig. 2B shows the cDNA encoding the *C. elegans* DAF-2.

Fig. 2C shows the amino acid comparison of *C. elegans* DAF-2 to the human insulin receptor and human IGF-I receptor (shown in parenthesis), and to the *Drosophila* insulin receptor homolog, with *daf-2* and human insulin receptor mutations highlighted. Six *daf-2* mutations map in the ligand-binding domain: *sa187* (C347S, TGT to AGT), *e1368* (S451L, TCA to TTA), *e1365* (A458T, GCT to ACT), *sa229* (D526N, GAT to AAT), and two mutations in *mg43* (C279Y, TGT to TAT and P348L, CCC to CTC). Three *daf-2* mutations substitute conserved amino acid residues in the insulin receptor kinase domain: *sa219* (D1252N, GAT to AAT), *e1391* (P1312L, CCC to CTC), and *e1370* (P1343S, CCA to TCA). Darkened residues indicate amino acid identity. Hatched residues indicate amino acid similarity. The percentages under the

domains represents the percentage of identity observed between DAF-2 and each receptor. The corresponding BLAST probabilities of DAF-2 random match to each protein is: 6.4×10^{-157} (human insulin receptor), 2.7×10^{-156} (human IGF-I receptor), 2.1×10^{-153} (molluscan InR homolog), 8.3×10^{-153} (mosquito InR homolgoue), 1.6×10^{-138} (human insulin receptor-related receptor), 1.7×10^{-122} (*Drosophila* InR homolog), 2.0×10^{-108} (Hydra InR homolog). DAF-2 is more distant from the next most closely related kinase families: 8.9×10^{-58} (v-ros) and 3.0×10^{-51} (trkC neurotrophin receptor).

Conserved cysteine residues in the ligand-binding domain (top) are marked with dots. In the kinase domain, active site residues that mediate insulin receptor kinase specificity are marked with stars. All of these residues are homologous in DAF-2. The mutations found in human patients are indicated at the top of the row, and *daf-2* allele substitutions are indicated below with allele names. The sequence alignments were done with GCG programs, Pileup and Prettybox, and the identities were calculated with the GCG program, Gap.

Fig. 3 is a photograph showing the metabolic control by *C. elegans* *daf-2* and *daf-7*. The top panel shows low levels of fat accumulation in a wild type L3 animal grown at 25°C that has been stained with Sudan black. Non-starved animals were fixed in 1% paraformaldehyde in PBS, frozen at -70°C, and freeze-thawed three times. Fixed animals were washed three times in PBS, and then incubated overnight in 1X Sudan black according to standard methods. The next panel shows higher levels of fat accumulation in *daf-2(e1370)* grown at the non-permissive temperature of 25°C. These animals accumulate fat in both intestinal and hypodermal cells. *daf-2(e1370)* animals grown at 15°C, the permissive temperature, accumulate low levels of fat, like wild type (data not shown). The next panel shows high fat levels in the intestine and hypodermis of *daf-7(e1372)* animals grown at 25°C. The bottom

panel shows high levels of fat in *daf-2(e1370)* animals grown at the permissive temperature until the L4 stage and then shifted to the non-permissive temperature. This shows that *daf-2* regulates metabolism without entry into the dauer stage.

Fig. 4 is a schematic diagram showing a model of insulin signaling in the *C. elegans* dauer formation pathway. In the absence of dauer pheromone, an insulin-like ligand activates DAF-2, and DAF-7 TGF-β-like signal activates the DAF-1 and DAF-4 receptors. Activated DAF-2 autophosphorylates particular tyrosine residues and recruits signaling molecules, including the PI 3-kinase homolog (a heterodimer of an as yet unidentified p85 homolog and the PI 3-kinase catalytic subunit AGE-1). The AGE-1 PI 3-kinase produces PIP3 second messenger. This second messenger may regulate glucose transport (White and Kahn, 1994 *supra*), metabolic kinase cascades that include AKT and GSK-3 (Hemmings, *Science* 226:1344-1345, 1984; Jonas et al., *Nature*, 385:343-346, 1997), and transcription and translation of metabolic genes (White and Kahn, 1994, *supra*). DAF-16 acts downstream of DAF-2 and AGE-1 in this pathway and is negatively regulated by them (Vowels and Thomas, *Genetics*, 130:105-123, 1992; Gottlieb and Ruvkun, *Genetics*, 137:107-110, 1994). While both the DAF-7/TGF-β and DAF-2/insulin signaling pathways converge to control dauer formation, only the DAF-2 pathway controls reproductive phase longevity. This may be due to non-transcriptional outputs of DAF-2 suggested by precedents from insulin receptor signaling. DAF-7 signaling output is predicted to be only transcriptional as described herein.

Fig. 5A shows that *C. elegans daf-3* was genetically mapped to a region on the X chromosome between *aex-3* and *unc-1*. Cosmid and plasmid clones from the region were assayed for transformation rescue (Mello et al., *EMBO J* 10: 3959-3970,1991). Plasmid pRF4 (*rol-6* transformation marker, 100 ng/ml),

and cosmids (5-6 ng/ml) were injected into the gonad of *daf-7* (*e1372*); *daf-3* (*e1376*) animals. Transgenic animals were scored for dauer formation at 25°C; a dauer (i.e., a return to the *daf-7* phenotype) indicates rescue of *daf-3*; clones that rescue *daf-3* are boxed. B0217 rescues the *daf-3* phenotype; eighteen of nineteen transgenic lines were rescued (~80% dauers). Examination of sequence provided by the *C. elegans* Sequencing Consortium revealed a Smad homologous gene on B0217. A 13 kb subclone of B0217 containing just the Smad also rescues *daf-3* (see Fig. 3). No rescue was seen upon injection of other cosmids from the region, B0504 (7 lines tested, <1% rescue) and C05H10 (10 lines tested, <1% rescue). *mgDf90* is a deletion that removes all of *daf-3*.

Fig. 5B shows the structure of the *C. elegans daf-3* coding region. The top is the exon/intron structure of *daf-3*; coding exons are filled boxes, non-coding regions are open boxes, and lines are introns. *daf-3* cDNAs were isolated according to standard methods. Four cDNAs were sequenced completely; their N-termini are indicated by vertical lines. These three cDNAs contain ~400 bp of 3' UTR, but no poly-A tail; a *C. elegans* consensus poly-adenylation sequence is found 12 bp from the 3' end of the cDNAs. The longest of this cDNA appears full-length, as it contains a methionine codon and the genomic sequence contains no other methionine codon and no putative splice sites upstream before in-frame stop codons. To further characterize the 5' end of *daf-3*, PCR products from libraries or individual *daf-3* cDNAs were sequenced. From DNA isolated from a cDNA library, we amplified a product with a primer to SL1 and to a region in conserved domain I (shown as primer 1). For the individual cDNAs, we amplified with a primer to the cDNA vector and primer 1. These PCR products were sequenced from primer 2 to the 5' end, and we found that there is alternative splicing at the 5' end of *daf-3*, upstream of the conserved domains. The two alternate splice forms are indicated, and the ends of individual cDNAs are indicated by vertical lines. Note that the second

has the trans-spliced leader SL1 that is found at the 5' end of many *C. elegans* cDNAs; thus, this cDNA shows a *bona fide* 5' end of *daf-3*.

Fig. 5C shows the protein sequence alignment of *C. elegans daf-3* and the closest homolog found to date, human DPC4, in the Smad conserved domains I and II. Dots indicate gaps introduced to maximize alignment. DAF-3 is 55% identical to DPC4 in domain I and 30% identical in domain II. *daf-3*(mg125) and *daf-3(mg132)* mutations are indicated by boldface and underline. The Smad mutational hotspot is underlined. In addition to *mg125* and *mg132*, seven other *daf-3* alleles were sequenced in the hotspot; none of them contains a mutation. Alleles sequenced were *mg91*, *mg93*, *mg105*, *mg121*, *mg126*, *mg133* (isolated by A. Koweek and G. Patterson, unpublished) and sa205.

Figs. 6A-6G is a panel of photographs showing *C. elegans* DAF-3 and DAF-4 expression. These photographs show GFP fluorescence, paired with DAPI fluorescence or Nomarski optics photographs, as marked. All DAF-3 photographs show animals with the second plasmid from **Fig. 6A** illustrates DAF-3/GFP head expression in an L1 animal. **Fig. 6B** illustrates DAF-3/GFP expression in the ventral nerve cord of an adult animal. L1 animals demonstrated similar expression patterns. **Fig. 6C** illustrates DAF-3/GFP expression in the intestine of an L1 animal. **Fig. 6D** illustrates DAF-3/GFP expression in the distal tip cell of an L4 animal. **Fig. 6E** illustrates DAF-3/GFP expression in an embryo with approximately 200 nuclei. **Fig. 6F** illustrates DAF-4/GFP expression in the head of an L1 animal. **Fig. 6G** illustrates DAF-4/GFP expression in the dorsal nerve cord and ventral nerve cord of an L4 animal.

Fig. 7 is a table that shows the rescuing ability and suppression of *C. elegans daf-7* by *daf-3* plasmids. The solid boxes represent the Smad conserved domains I and II of *daf-3*; the stippled boxes represent green fluorescent protein (GFP). For all experiments shown, *daf-3* plasmids were

injected at a concentration of 10 ng/ml, and the pRF4 injection marker was injected at a concentration of 90 ng/ml. To score dauer formation, transgenic adult animals were allowed to lay eggs on plates for several hours at room temperature and were then removed. The plates were scored after two days at 25°C. The rescue experiment shows the rescue of *daf-7(m62); daf-3(e1376)* by each of the fusion proteins. Failure to rescue results in rolling nondauers, while rescue of *daf-3* results in rolling dauers (the *daf-7* phenotype). The control is an array with the pRF4 transformation marker and a non-rescuing cosmid. For each construct, four or more lines were measured in two separate experiments. To measure suppression of *daf-7*, transgenic arrays were crossed into *daf-7* (for plasmids 1 and 3), or produced by injecting directly into *daf-7* (for plasmid 2). Transgenic (rolling) animals were scored for suppression of *daf-7* (= nondauers) or failure to suppress *daf-7* (= dauers). The controls are two array strains with the pRF4 marker and an unrelated GFP expressing transgene.

Fig. 8A is a photographs showing that DAF-3/GFP is associated with metaphase chromosomes. Fixed L1 animals were immunostained with anti-GFP antibody and anti-α-tublin antibody. DNA was visualized using DAPI staining.

Fig. 8B is a photograph showing that a truncated *C. elegans daf-3*/GFP protein is predominantly nuclear. Wild-type animals were injected with the truncated construct shown in **Fig. 7** at a concentration of 10 ng/ml. The pRF4 transformation marker was injected at 100 ng/ml. The photograph shows a late L1 or early L2 animal, and *daf-3* is predominantly nuclear. The clear spot in the center of some of the nuclei is the nucleolus, which has no *daf-3*/GFP. All cells in these animals have predominantly nuclear *daf-3*/GFP, including the ventral cord neurons, intestinal cells, and distal tip cell (all shown), as well as head and tail neurons and hypodermal cells.

Figs. 9A and 9B show models for the role of the *C. elegans daf-3*/DAF-8/DAF-14 Smad proteins in dauer formation. **Fig. 9A** shows dauer reproductive growth induction. **Fig. 9B** shows reproductive dauer growth induction.

Fig. 10 is a schematic illustration showing the genetic pathway that regulates *C. elegans* dauer formation.

Figs. 11A-11C show the cDNA sequences of the differentially spliced *C. elegans daf-3* transcripts (SEQ ID NOS: 39, 52, and 53).

Figs. 12A-12C show the amino acid sequences of the *C. elegans* DAF-3 polypeptide isoforms (SEQ ID NOS: 40-42).

Figs. 13A and 13B show the cDNA sequence of the differentially spliced *C. elegans daf-16* transcripts (SEQ ID NOS: 43 and 44).

Figs. 14A and 14B show the amino acid sequences of the C. elegans DAF-16 polypeptide isoforms (SEQ ID NOS: 45 and 46).

Fig. 15 shows the cDNA sequence of the *C. elegans age-1* gene (SEQ ID NO: 47).

Fig. 16 shows the amino acid sequence of the *C. elegans* AGE-1 polypeptide (SEQ ID NO: 48).

Fig. 17 is a schematic diagram illustrating that convergent TGF-β and insulin signaling activates glucose-based metabolic genes.

Fig. 18 is a schematic diagram illustrating a switch to fat-based metabolism in the absence of DAF-7 and DAF-2 signals (in phermone).

Fig. 19 is a schematic diagram illustrating inhibition of the DAF-16 pathway by drugs to ameliorate lack of insulin signaling.

Fig. 20 is a schematic diagram illustrating inhibition of DAF-3 by drugs to ameliorate a lack of DAF-7 signaling (for example in obesity-induced diabetes).

Fig. 21A is an illustration showing that human FKHR and AFX are the

closest relatives to DAF-16. Note that the differentially spliced DAF-16 forkhead domain is less homologous.

Fig. 21B is an illustration showing a forkhead family tree, illustrating that DAF-16 is much more closely related to FKHR and AFX than any other forkhead protein.

Fig. 22 is a photograph showing that *daf-16* is expressed in target tissues, like *daf-3*. This supports the model that DAF-3 and DAF-16 are capable of interacting.

Fig. 23 is an illustration showing a model for treatment of obesity-induced diabetes with DAF-7 protein.

Fig. 24 is an illustration showing the genetic mapping of *sup(mg144)* to the AKT genetic region.

Fig. 25 is an illustration showing the comparison of *C. elegans* AKT with mammalian AKT.

Fig. 26A is a photograph showing the expression of AKT:GFP in *daf-2* dauers.

Fig. 26B is a photograph showing the expression of AKT:GFP in an N2 adult worm.

Fig. 27 is a schematic illustration showing the molecular map of *daf-16*.

Fig. 28 is a graph illustrating the homology of *C. elegans* insulin-like molecules (SEQ ID NOS: 117-124) with human insulin (SEQ ID NO: 125) and a consensus motif.

Fig. 29 is a graph illustrating a PRETTYBOX analysis of insulin superfamily members (SEQ ID NOS: 126-153).

Fig. 30 is a graph illustrating a PILEUP analysis of insulin superfamily members.

Fig. 31 is a diagram illustrating the *akt-1* region. On the top is shown the genetic and physical map of *akt-1*. *akt-1* is contained on cosmid C12D8.

Shown on the bottom is the exon/intron structure of *akt-1*. Coding regions are filled boxes, non-coding regions are open boxes, and introns are lines. The pleckstrin homology domain is indicated by hatched boxes (Musacchio et al., Trends Biochem. Sci. 18:343-348, 1993). The kinase domain is indicated in gray (Hanks and Hunter, in The Protein Kinase Facts Book Protein-Serine Kinases, eds. Hardie, G. & Hanks, S., Academic Press, Inc., San Diego, CA, pp. 7-47, 1995). *akt-1a* gene structure was confirmed by sequencing of cDNAs. *akt-1b* gene structure was deduced based on partial cDNA sequence that confirmed the exon 5 to exon 7 splice and 3'UTR only.

Fig. 32 is a diagram illustrating the *akt-2* region. On the top is shown the genetic and physical maps of the *akt-2* region. *akt-2* is contained on cosmid R03E1. On the bottom is shown the exon/intron structure of *akt-2*. All symbols are as in **Fig. 31**. Gene structure was deduced by sequencing of a cDNA which confirmed exons 2-8 and the 3'UTR; Genefinder (Univ. of WA) predicts exon 1.

Fig. 33 is a graph illustrating a dendogram of Akt/PKB and PKC protein kinase families. Pileup (GCG) was used to align the entire coding sequences of the indicated proteins. *C. elegans* proteins are indicated by "Ce," rat by "r," human by "h," mouse by "m," bovine by "b," and *D. melanogaster* by "D." The accession numbers for the proteins used in the Pileup are contained in parentheses: CePKC2a(U82935), rPKCβ1(M19007), hAkt/PKBα(M63167), mAkt/PKB(M94335), bAkt/PKB(X61036), hAkt/PKBβ2(M95936), rAkt/PKBγ(D49836), Dakt1(Z26242). To anchor the tree, rPKCβ1 (the closest non-Akt/PKB homolog to both *akt-1a* and hAkt/PKBα), and CePKC2a (the closest *C. elegans* homolog to rPKCβ1) were included in the Pileup. The Akt/PKB homologs described in this report are indicated by the gray box.

Fig. 34 is a graph illustrating a PILEUP (GCG) analysis of AKT-1a (SEQ ID NO: 154), AKT-1b (SEQ ID NO: 155), AKT-2 (SEQ ID NO: 156),

and human Akt/PKBα (M63167) (SEQ ID NO: 157). Identical residues are indicated by dots, gaps introduced in order to align the sequence are indicated by dashes. The pleckstrin homology domain (Musacchio et al., Trends Biochem. Sci. 18:343-348, 1993) is indicated by the N-terminal gray shaded areas, the kinase domain (Hanks and Hunter, in The Protein Kinase Facts Book Protein-Serine Kinases, eds. Hardie, G. & Hanks, S., Academic Press, Inc., San Diego, CA, pp. 7-47, 1995) is indicated by the C-terminal gray shaded areas. The mg144 Ala183Thr substitution is indicated as a T above the AKT-1a sequence. The Akt-1 and AKT-2 phosphorylation sites that correspond to the hAkt/PKBα Thr308 and Ser473 phosphorylation sites (Alessi et al., EMBO J. 15:6541-6551, 1996) are indicated as dots above the amino acid residue that is phosphorylated.

Figs. 35A and 35B show the genomic sequence of *pdk-1* (SEQ ID NO: 158).

Fig. 36 shows the amino acid sequence of *pdk-1a* (SEQ ID NO: 159).

Fig. 37 shows the amino acid sequence of *pdk-1b* (SEQ ID NO: 160).

The DAF-2 Insulin Receptor Family Member Regulates Longevity and Diapause in C. elegans

Arrest at the *C. elegans* dauer stage is normally triggered by a dauer-inducing pheromone detected by sensory neurons which signal via a complex pathway to target tissues that are remodeled and metabolically shifted such as the germ line, intestine, and ectoderm (Riddle, *In*: *Caenorhabditis elegans* II, D. Riddle, T. Blumenthal, B. Meyer, J. Priess, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1997, pp. 739-768. Kenyon, op cit., pp. 791-813.). Genetic epistasis analysis of *daf* mutants that arrest at the dauer stage or enter the reproductive life cycle independent of pheromone regulation has revealed parallel genetic pathways that regulate distinct aspects

of the dauer metamorphosis (Vowels and Thomas, *Genetics* 130: 105-123, 1992; Gottlieb and Ruvkun, *Genetics* 137: 107-120, 1994). The pathway that includes *daf-2* is unique in that it controls both reproductive development and normal senescence: *daf-2* mutant animals arrest development at the dauer larval stage and have dramatically increased longevity (Table I) (Riddle, In: *Caenorhabditis elegans* II, D. Riddle, T. Blumenthal, B. Meyer, J. Priess, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1997, pp. 739-768; Kenyon, op cit. pp 791-813; Vowels and Thomas, Genetics 130: 105-123, 1992; Gottlieb and Ruvkun, Genetics 137: 107-120, 1994; Larsen et al., Genetics 139: 1567-1583, 1995; Kenyon et al., Nature 366: 461-464, 1993; Dorman et al., Genetics 141: 1399-1406, 1995).

Table I shows the percentage of dauer formation of *daf-2* alleles and the associated mutations. Eggs from animals grown at 15°C (day 0) were incubated at 15, 20, or 25°C. Numbers in parenthesis are animals counted. Numbers of wild-type animals and dauers were counted on day 3 (20°C and 25°C) or day 5 (15°C). Most of the dauers marked with stars recovered by day 4 (*sa229* at 25°C) or by day 8 (*sa229*) and *sa219* at 15°C, *e1368* and *sg219* at 20°C, and *e1365* and *e1368* at 25°C). *mg43* was studied as follows: *dpy-1(e1)daf-2(mg43); SDP3* animals were grown at 20°C until the young adult stage. Eggs from five adults were laid at 15°C or 20°C and grown at the same temperatures. Numbers of Dpy-Daf animal and Dpy-non-Daf animals were counted on day 3 (20°C) or day 5 (15°C). *Sg187* and *sg229* were also studied by Malone and Thomas (*Genetics* 136:879-886, 1994).

Table I. Percentage of dauer formation of *daf-2* alleles

Region	Allele	mutation	% dauer formation 15°C	20°C	25°C
cys-rich	mg43	C279Y&P348L	100.0 (215)	100.0 (245)	n.d.
	sa187	C347S	0.4 (461)	98.7 (224)	100 (910)
ligand-binding	e1368	S451L	0.0 (328)	4.5* (418)	99.7* (698)
	e1365	A458T	0.0 (450)	0.0 (461)	99.4* (814)
	sa229	D526N	3.4* (234)	n.d.	22.1* (420)
	sa219	D1252N	10.0* (460)	99.7* (396)	100 (514)
kinase	e1391	P1312L	3.3 (332)	100 (323)	100 (322)
	e1370	P1343S	0.0 (520)	0.0 (188)	100 (635)

Genetic mapping using both visible genetic markers and restriction fragment length polymorphism (RFLP) markers places *daf-2* between *mgP34* and *mgP44* (Fig. 1). While cosmid coverage of this physical genetic region is not complete, YAC Y53G8 carries the genomic region that includes *mgP34* and *mgP44*, which flank *daf-2* (Fig. 1). As a step in the *C. elegans* genome sequencing effort, random M13 subclones derived from Y53G8 were sequenced by the Genome Sequencing Center.

Sequence Identities Show that DAF-2 is Likely to Bind to an Insulin-like Ligand and to Phoshorylate Tyrosine Residues

The amino acid sequences and nucleotide sequences encoding DAF-2 are shown in Figs. 2A and 2B, respectively. Using BLASTX to compare 570 translated Y53G8 M13 subclone sequences against the Genbank protein database, we found that four sequences are homologous to the mammalian insulin receptor family. An insulin receptor was a good *daf-2* candidate gene because insulin regulates vertebrate growth and metabolism (White and Kahn,

J. Biol. Chem. 269: 1-4, 1994), and because a phosphatidylinositol (PI) 3-kinase has been shown to act in both the insulin receptor and *daf-2* pathways (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994; Morris et al., *Nature* 382: 536-539, 1996). The detection of multiple *daf-2* mutations in the gene (see below), and the coincidence of the genetic location of this insulin receptor homolog with *daf-2* (Fig. 2C) establish that this insulin receptor homolog corresponds to *daf-2*.

The *daf-2* transcription unit and gene structure were determined using PCR primers derived from *daf-2* genomic subclone sequences to amplify *daf-2* genomic and cDNA regions. A probable full length *daf-2* cDNA bears a 5172 base open reading frame, a 485 base 5' UTR and a 159 base 3' UTR (Figs. 1, 2A). The predicted DAF-2 protein shows long regions of sequence identity to the insulin receptor family. Over the entire protein, DAF-2 is 35% identical to the human insulin receptor (Ebina et al., *Cell* 40: 747-58, 1985; Ullrich, et al., *Nature* 313: 756-61, 1985), 34% identical to the human IGF-1 receptor (Ullrich, et al., *EMBO J.*: 5, 2503-12, 1986), and 33% identical to the human insulin receptor-related receptor (Shier and Watt, *J. Biol. Chem.* 264: 14605-8, 1989). DAF-2 is the only member of the insulin receptor family in the 90 Mb *C. elegans* genome sequence (about 90% complete) or in the 10 Mb *C. elegans* EST sequence database. Because it is equally distant from insulin, IGF-1, and insulin receptor-related receptors, DAF-2 is probably the homolog of the ancestor of these duplicated and diverged receptors, and thus may subserve any or all of the functions of these mammalian receptors (see below). Like these receptors, DAF-2 has a putative signal peptide, a cysteine-rich region in the putative ligand binding domain, a putative proteolysis site, a transmembrane domain, and a tyrosine kinase domain. In addition, DAF-2 has a C-terminal region that may serve a function similar to the mammalian insulin receptor substrate-1 (IRS-1) (Figure 2; White and Kahn,

J. Biol. Chem. 269: 1-4, 1994).

In the approximately 500 amino acid ligand-binding domain of the insulin receptor, DAF-2 is 36% identical to insulin receptor and 35% identical to the IGF-I receptor. Twenty-one of twenty-three phylogenetically conserved cysteine residues in this domain are conserved in DAF-2 (Fig. 2C). The DAF-2 cys-rich region is 34% identical to human insulin receptor and 28% identical to the IGF-I receptor. Six *daf-2* mutations map in this domain (Fig. 2C, Table I). The *mg43* and *sa187* mutations substitute conserved residues in the cys-rich region (Fig. 2C). *daf-2(mg43)* carries two mutations which substitute conserved residues, which may explain the strength of this allele (non-conditional, Table I). Other substitutions at non-conserved residues cause less severe phenotypes (Table I). Insulin resistant and diabetic patients with mutations in the ligand binding domain of the human insulin receptor gene have been identified (Taylor, *Diabetes* 41: 1473-1490, 1992) (see below). These mutations impair receptor transport to cell surface, or insulin binding affinity, or both. The DAF-2 mutations in this domain might similarly decrease receptor signaling to cause dauer arrest.

Insulin receptors are $\alpha 2,\beta 2$ tetramers proteolytically processed from a single precursor protein (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). DAF-2 bears a probable protease recognition site at a position analogous to the insulin receptor processing site (RVRR 806-809) (Yoshimasa et al., *J. Biol. Chem.* 265: 17230-17237, 1990).

The 275 amino acid DAF-2 tyrosine kinase domain is 70% similar and 50% identical to the human insulin receptor kinase domain. Upon insulin binding, the intracellular tyrosine kinase domain of the insulin receptor phosphorylates particular tyrosine residues flanked by signature amino acid residues (upstream acidic and downstream hydrophobic amino acids (Songyang and Cantley, *Trends Biochem. Sci.* 20: 470-475, 1995)) in the intracellular

domain as well as on IRS-1 (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). Multiple DAF-2 tyrosine residues in these sequence contexts are likely autophosphorylation targets, including three tyrosines in a region similar to the insulin receptor activation loop and one in the juxtamembrane region as described above (Fig. 2C). Based on the crystal structure of the insulin receptor kinase domain bound to its activation loop, eight kinase domain residues mediate target site specificity (Hubbard et al., *Nature* 372: 746-754, 1994). In DAF-2 (but not in more distantly related receptor kinases), these residues are invariant (5/8) or replaced with similar amino acids (3/8: K to R, E to D) (Fig. 2C), suggesting that DAF-2 phosphorylates the same target tyrosine motifs as the insulin receptor kinase.

Three *daf-2* missense mutations substitute conserved amino acid residues in the kinase domain (Fig. 2C, Table I). All three mutations cause moderate to strong dauer constitutive phenotype, but none are as strong as the non-conditional alleles, for example, *mg43* (Table I). Human insulin receptor mutations in the kinase domain exhibit decreased kinase activity and cause severe insulin resistance and associated defects (Fig. 2C; Taylor, *Diabetes* 41: 1473-1490, 1992). Remarkably, a human diabetic insulin resistant patient bears the same amino acid substitution (P1178L) as *daf-2(e1391)* (Kim et al., *Diabetologia* 35: 261-266, 1992). This patient was reported to be heterozygous for this substitution. *daf-2(e1391)* is not dominant whereas it is a highly penetrance recessive mutation (Table I).

To test for dominance of *daf-2(e1391)*, using a genetically marked balancer chromosome, 105 dauers segregated from 485 daf-2/+ parents as expected for a recessive mutations. The genotype of 76/77 of these animals was homozygous *daf-2(e1391)* whereas 1/77 of the dauers was *daf-2(e1391)*/+, indicating a less than 1% dominance. It is possible that in contrast to *C. elegans*, the P1178L mutation in humans is dominant, or that the patient carries

a second insulin receptor mutation in *trans*, or carries mutations in other genes (for example, other complex type II diabetes loci) that enhance the dominance of P1178L (Bruning et al., *Cell* 88: 561-572, 1997).

AGE-1 PI 3-kinase is a Major DAF-2 Signaling Output

Like the *Drosophila* insulin receptor homolog, DAF-2 has a long C-terminal extension that may function analogously to mammalian IRS-1 (Fernandez et al., *EMBO* J. 14: 3373-3384, 1995). In mammals, IRS-1 tyrosine residues are phosphorylated by the insulin receptor kinase, and these phosphotyrosines mediate binding to a variety of signaling proteins bearing SH2 domains (White and Kahn, J. *Biol. Chem.* 269: 1-4, 1994; Songyang et al., *Cell* 72: 767-778, 1993.). Many, but not all, of the DAF-2 C-terminal extension tyrosines bear flanking sequence motifs suggestive that they are autophosphorylated (Fig. 2A; Songyang and Cantley, *Trends Biochem.* Sci. 20: 470-475, 1995). Based on precedents from IRS-1 interactions with mammalian PI 3-kinases (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994), a YXXM motif at DAF-2 Y1504 is likely to mediate interaction with the AGE-1 PI 3-kinase, which acts in the same genetic pathway as *daf-2* (Fig. 4) (Morris et al., *Nature* 382: 536-539, 1996).

Three DAF-2 tyrosine residues, Y1293, Y1296 and Y1297, in the region corresponding to the insulin receptor kinase Y1158 to Y1163 activation loop are likely to be autophosphorylated, based on the predicted similarity between the DAF-2 and insulin receptor phosphorylation targets (Fig. 2C). Another likely target for DAF-2 autophosphorylation is the Y1106 NPEY motif located in the region corresponding to the insulin receptor juxtamembrane region NPEY motif (at Y972), that has been shown to mediate IRS-1 binding via its PTB domain to the insulin receptor (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). While DAF-2 bears one YXXM motif implicated in coupling to PI

3-kinase, mammalian IRS-1 and *Drosophila* insulin receptor (Fernandez et al., *EMBO J.* 14: 3373-3384, 1995) bear multiple YXXM motifs. Although no p85-like adaptor subunit has yet been detected in the *C. elegans* database, the AGE-1 homology to mammalian p110 suggests the existence of a homologous or analogous adaptor (Morris et al., *Nature* 382: 536-539, 1996). In the DAF-2 C-terminal domain, two other tyrosine residues may be autophosphorylated and bound to particular SH2-containing proteins: Y1678 binding to a PLC-γ or SHP-2 homolog, and Y1686, perhaps binding to SEM-5 (Fig. 2A) (Songyang et al., *Cell* 72: 767-778, 1993). While mutations in, for example, ras and MAP kinase have not been identified in screens for dauer constitutive or dauer defective mutations, these general signaling pathway proteins may couple to DAF-2 as they couple to insulin signaling in vertebrates (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994).

The insulin receptor also couples to other signaling pathways (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994); analogous DAF-2 phosphotyrosine residues may mediate these interactions (as described above). Thus, we suggest that tyrosines in the DAF-2 cytoplasmic domain are autophosphorylated upon ligand binding, and recruit the AGE-1 PI-3 kinase homolog (as well as other molecules) to signal reproductive development and normal senescence.

Metabolic Control by daf-2 in Control of Diapause and Aging

Insulin and its receptor families play key roles in vertebrate (and by our evidence in invertebrates) metabolic and growth control (Kahn and Weir, eds., *Joslin's Diabetes Mellitus*, Lea & Febiger, 1994). Upon insulin release--by increasing blood glucose and autonomic inputs--insulin receptor engagement directs a shift in the activities of key metabolic enzymes, as well as changes in the transcription and translation of metabolic regulators in fat, liver, and muscle

cells, all of which lead to assimilation of glucose into glycogen and fat (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). IGF-I is released from the liver in response to pituitary growth hormone, and mediates many of the growth and development responses to that endocrine signal (Mathews et al., *Proc Natl Acad Sci. U.S.A.* 83: 9343-7, 1986). Interestingly, lifespan is dramatically increased in dwarf mice with defects in growth hormone signaling, and presumably decreased IGF-I signaling as well (Brown-Borg et al., *Nature* 384: 33, 1996). No function for the insulin receptor-related receptor has yet been established, though it is expressed in conjunction with NGF receptor (Reinhardt et al., *J. Neurosci.* 14: 4674-4683, 1994).

Diapause arrest in general and dauer arrest in particular are associated with major metabolic changes (Tauber et al., *Seasonal Adaptation of Insects*, Oxford University Press, New York, N. Y., 1986), consistent with a model that *daf-2* acts in a metabolic regulatory pathway related to insulin signaling. In wild-type animals, DAF-2 signaling allows non-dauer reproductive growth, which is associated with utilization of food for growth in cell number and size, and small stores of fat (Fig. 3). In *daf-2* mutant animals, metabolism is shifted to the production of fat (Fig. 3) and glycogen (data not shown) in intestinal and hypodermal cells. Even when a temperature-sensitive *daf-2* mutant allele is shifted to the non-permissive temperature at the L4 or adult stage (after the critical period for *daf-2* control of dauer formation), metabolism is shifted towards storage of fat (Fig. 3). Thus *daf-2* also regulates metabolism during reproductive development. Similar metabolic shifts are seen in wild-type pheromone-induced dauers (data not shown), *age-1* mutants (data not shown), and *daf-7* mutants (Fig. 3). In support of this metabolic shift, in dauer larvae, enzymes that regulate glycolysis are down-regulated while those that regulate glycogen and fat synthesis are up-regulated, and there is ultrastructural evidence for increased lipid and glycogen (O'Riordan and Burnell, *Comp.*

Biochem. & Physiol. 92B: 233-238, 1989; O'Riordan and Burnell, *Comp. Biochem. & Physiol.* 95B: 125-130, 1990; Popham and Webster, *Can. J. Zool.* 57: 794-800, 1978; Wadsworth and Riddle, *Develop. Biol.* 132: 167-173, 1989). The dauer metabolic shift is associated with arrest of germ line proliferation, and arrest of somatic cell division and enlargement (Riddle, In: *Caenorhabditis elegans* II, D. Riddle, T. Blumenthal, B. Meyer, J. Priess, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1997, pp. 739-768; Kenyon, op cit., pp. 791-813).

There is precedent for insulin-like signaling in invertebrate metabolic and growth control: insulin-like growth factors have been detected in metabolism-regulating ganglia in molluscs (Roovers et al., *Gene* 162: 181-188, 1995) and regulate molting in locust (Hetru et al., *Eur. J. Biochem* 201: 495-499, 1991) and silkworm (Kawakami et al., *Science* 247: 1333-1335, 1990). Consistent with the *daf-2* regulation of diapause, injection of insulin into diapausing *Pieris brassicae* (an insect) pupae induces recovery (Arpagaus, *Roux's Arch. Dev. Biol.* 196: 527-530, 1987).

Without being bound to a particular theory, we hypothesize that an insulin-like signal is up-regulated during reproductive development and stimulates DAF-2 receptor autophosphorylation and recruitment of the AGE-1 PI 3-kinase to produce the second messenger PIP3. AGE-1 is likely to be a major signaling output of DAF-2 because of the similarity of the *age-1* and *daf-2* mutant phenotypes and because of their similar placement in the epistasis pathway (Vowels and Thomas, *Genetics* 130: 105-123, 1992; Gottlieb and Ruvkun, *Genetics* 137: 107-120, 1994). Precedents from insulin receptor signaling suggest the following candidate targets for DAF-2/AGE-1/PIP3 regulation of metabolism: (1) membrane fusion of vesicles bearing glucose transporters (Kahn and Weir, eds., *Joslin's Diabetes Mellitus*, Lea & Febiger, 1994) (or more probably trehalose transporters (Tauber et al., *Seasonal*

Adaptation of Insects, Oxford University Press, New York, N. Y., 1986)) to facilitate flux of this molecule for growth and reproductive metabolism; (2) PIP3 activates an AKT/GSK-3 kinase cascade (Hemmings, *Science* 275: 628-630, 1997) which may regulate the activities of glycogen and fat synthetic and lytic enzymes; (3) transcription and translation of metabolic genes such as PEPCK, GDH, fat synthetases, and lipases (White and Kahn, *J. Biol. Chem.* 269:1-4, 1994). Genetic epistasis analysis suggests that DAF-2/AGE-1 signaling negatively regulates *daf-16* gene activity (Vowels and Thomas, *Genetics* 130: 105-123, 1992; Gottlieb and Ruvkun, Genetics 137: 107-120, 1994). DAF-16 could act at any point downstream of AGE-1 in this signaling pathway. Evidence is presented herein that DAF-16 represents the major transcriptional output to DAF-2/AGE-1 PIP3 signaling.

In addition to these metabolic changes, the DAF-2 signaling cascade also controls the reproductive maturation of the germ line as well as morphogenetic aspects of the pharynx and hypodermis (Riddle, In: *Caenorhabditis elegans* II, D. Riddle, T. Blumenthal, B. Meyer, J. Priess, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1997, pp. 739-768; Kenyon, op cit., pp. 791-813). The DAF-2 receptor may act, for example, in the hypodermal and intestinal target tissues where we note a change in metabolism triggered by the dauer regulatory cascade (Fig. 3). It is also possible that DAF-2 regulates the metabolism and remodeling of tissues indirectly, for example, by controlling the production of other hormones (Nagasawa et al., *Science* 226: 1344-1345, 1984; Jonas, et al., *Nature* 385: 343-346, 1997). Expression and genetic mosaic analysis of *daf-2* is essential to distinguish these models.

Even though DAF-2 and the mammalian insulin receptor both regulate metabolism, the metabolic defects associated with mutations in these receptors appear to be different. Complete loss of mammalian insulin receptor activity

causes growth arrest at birth (Leprechaunism in humans), and a metabolic shift to runaway lipolysis and ketoacidosis (Kahn and Weir, eds., *Joslin's Diabetes Mellitus*, Lea & Febiger, 1994), rather than the fat accumulation we observe in *daf-2* mutants (Fig. 3). This distinction between insulin receptor and *daf-2* mutants may reflect distinct metabolic responses to this signaling, or a difference between complete loss and declines in insulin signaling. In humans, ketoacidosis is only induced during severe starvation or pathological states when insulin levels are very low (Kahn and Weir, eds., *Joslin's Diabetes Mellitus*, Lea & Febiger, 1994). Since none of the *daf-2* mutations described herein are clear null mutations, it is possible that *daf-2* dauer-constitutive alleles are more analogous to non-null human insulin receptor mutations. Most *daf-2* alleles are temperature sensitive, including alleles isolated in genetic screens that would allow the recovery of non-temperature sensitive mutations (Vowels and Thomas, *Genetics* 130: 105-123,1992; Gottlieb and Ruvkun, *Genetics* 137: 107-120, 1994). Substitutions of DAF-2 amino acid residues conserved across phylogeny cause more penetrant dauer arrest at all temperatures than substitutions of non-conserved residues. *daf-2* mutants that arrest development at the dauer stage independent of growth temperature are likely to have the least gene activity (for example *mg43*). Several *daf-2* alleles also cause 5% to 10% embryonic lethality (unpublished results), suggesting that *daf-2* functions during embryonic development. None of the *daf-2* mutations detected so far are nonsense, frameshift, or deletion alleles. It is possible that the *daf-2* null phenotype is stronger than non-conditional dauer arrest, for example embryonic lethality. However, dauer constitutive *daf-2* mutant alleles are isolated from EMS mutagenesis at a very high rate (about 1/300 chromosomes), suggesting that the existing alleles are not rare viable alleles. In fact, the 14 year old patient with the same insulin receptor mutation as *daf-2(e1391)* was morbidly obese (Kim et al., *Diabetologia* 35: 261-266,

1992), suggesting that metabolic effects of decreased insulin signaling may be similar to *daf-2* mutants.

It may be significant to human diabetes that animals carrying mutations in *daf-16* can grow reproductively even if they also carry *daf-2* and *age-1* mutations that disable insulin-like metabolic control signals (Vowels and Thomas, *Genetics* 130: 105-123, 1992; Gottlieb and Ruvkun, Genetics 137: 107-120, 1994). These data suggest that it is unregulated *daf-16* gene activity that causes these metabolic shifts. The analogous metabolic defects associated with both type I and type II diabetes may be caused by similar unregulated activity of the human DAF-16 homolog. Below we disclose the molecular identity of *daf-16*. Inhibition of its activity is expected to ameliorate the metabolic dysregulation associated with insulin signaling defects.

DAF-16 Encodes a Forkhead Transcription Factor Homolog

Using a combination of genetic mapping and detection of multiple *daf-16* mutations in a 5 kb region, we have determined the nucleic acid sequence of *daf-16*. *daf-16* was mapped 1 map unit to the left of *lin-11* and 3.3 map units right of *unc-75* on Chromosome I. This region of the genome contained a gap that was not covered by cosmids nor YACs. We used a fosmid library (Genome Sciences, Inc.) to walk into the gap. Sequence analysis of the ends of four fosmids (H27K20, H01H03, H12I08, and H35K06) revealed that the previously unmapped contig 133 lies in the *lin-11 unc-75* gap. Cosmids from the approximate daf-16 genetic location were used to detect RFLPs between *C. elegans* strains Bristol N2 and Bergerac RC301: *mgP45* on cosmid C39H11, *mgP46* on cosmid F28D9, *mgP49* on cosmid C35E7, *mgP50* is on cosmid C43H8. Zero out of 30 daf non-Unc recombinants carry the RC301 alleles of *mgP45* and *mgP50*. Two out of 30 Daf non-Unc recombinants carry the RC301 allele of *mgP49*. 10 out of 30 Daf non-Unc recombinants carry the

RC301 allele of *mgP46*. 1 out of 4 non-Lin Daf recombinants carry the N2 allele of *mgP45*. 4 out of 4 non-Lin Daf recombinants carry the N2 allele of *mgP49*. These data indicate that *daf-16* lies between cosmids C43H8 and C35E7. The *daf-16* gene was identified by identifying deletions (mgDf50) and point mutations (*mg53* and *mg54*) within the forkhead gene on the cosmid R13H8 (Fig. 27). There are two major *daf-16* transcripts whose sequences are shown in Fig. 13A and Fig. 13B (SEQ ID NOS: 43 and 44, respectively). The amino acid sequences coding for the DAF-16 isoforms are shown in Figs. 14A-14C (SEQ ID NOS: 44-46).

We have detected three daf-16 mutations: (1) a large deletion of conserved regions in *daf-16* (mg ΔF50) that proves that the *daf-16* null phenotype is a suppression of *daf-2* mutations; (2) a S to L substitution in exon 6 in daf-16 (mg 53) that alters a conserved WKNSIRH motif; and (3) a nonsense mutation in exon 3 in *daf-16* (*mg 54*) that is predicted to truncate one of the *daf-16* differentially spliced isoforms. Interestingly, this spliced isoform has a distinct forkhead DNA binding domain and is therefore expected to bind to distinct promoters or combinatorial partners. This mutant is a weak suppressor of *daf-2*, suggesting that both DAF-16 isoforms are necessary for metabolic control.

Sequence analysis has revealed that DAF-16 is a member of the forkhead (FH) transcription factor family (Figs. 21A-21B). This strong amino acid homology indicates that DAF-16 is a transcription factor. Our genetic analysis indicates that DAF-16 activity is regulated by the DAF-2/AGE-1 insulin signaling pathway. Precedent from another receptor kinase signaling pathway endorses this model: the *C. elegans* LIN-31 forkhead protein has been shown to be regulated by a tyrosine kinase signaling cascade from the LET-23 EGF receptor homolog (Kim, *Genes Dev.* 7: 933-947, 1993). Consistent with a model that DAF-16 acts downstream of insulin signaling, forkhead

transcription factors have also been implicated in metabolic regulation: another FH family member is mammalian HNF-3, an endoderm-specific transcription factor that acts at the same metabolic control protein promoters as HNF-1 and HNF-4, both of which are mutant in maturity onset diabetes of the young (MODY) (Yamagata et al., *Nature* 384: 455-458, 1996; Yamagata et al., *Nature* 384: 458-460, 1996).

The identification of DAF-16 as a forkhead transcription factor also explains much of the complex daf genetics of *C. elegans*. The convergence of DAF-7 TGF-β-like signaling and DAF-2 insulin-like signaling is also explained by our discovery that DAF-16 is a FH protein and DAF-3 is a Smad protein:

Precedent for an interaction between Smad and forkhead proteins has been found in *Xenopus*. Response to the TGF-β superfamily relative activin in early frog development is mediated by an interaction between the distant relative of DAF-16 called FAST-1, and the Smad protein, Smad2 (*Nature* 383: 600-608, 1996). These proteins bind to an enhancer element that is very similar to the myosin II promoter to which DAF-3 binds (see below). Thus our molecular and genetic data indicate that the DAF Smad proteins and DAF-16 FH protein interact on metabolic control promoters.

Interestingly, analogously to *daf-16* bypass of the need for DAF-2 insulin receptor signaling in *daf-16* mutant animals, *lin-31* mutations suppress the need for LET-23 EGF signaling in *C. elegans* vulval development. These findings indicate that the DAF-2 receptor, a downstream signaling molecule (AGE-1), and a transcription factor target DAF-16 are involved in insulin-like signaling in *C. elegans* development. Without being bound by any particular theory, we hypothesize that *C. elegans* insulin signaling via DAF-2 and AGE-1 activate DAF-16 transcriptional activity, so that in a *daf-2* or *age-1* mutant, or in dauer pheromone, DAF-16 acts as a repressor protein causing a metabolic

shift to fat metabolism. Our analysis of *daf-16* expression shows that, like DAF-3, it is expressed in target tissues (Fig. 22). Our evidence indicates that Smad protein transcription factors (e.g., DAF 3, DAF8, DAF14) and DAF-16 act on a common set of promoters as combinatorial transcriptional regulators. Thus, it is at these metabolic genes that DAF-7 and TGF-β-like and DAF-2 insulin-like signals converge to control metabolism. In addition, our evidence indicates that in the presence of DAF-2 signaling (mimicking high insulin), DAF-16 acts as an activator of transcription, causing a shift in metabolism toward glucose utilization for cell growth. The molecular analysis described herein suggests that lack of *daf-16* gene activity completely bypasses the need for insulin signaling in metabolic control by releasing metabolic control from DAF-16 repression. These data suggest that if a human DAF-16 homolog acts downstream of insulin signaling in humans, drugs could be developed that inhibit its activity to bypass the need for insulin signaling. Identification of a such a drug should provide a means for treating both Type I and Type II diabetes.

As shown in Figs. 21A-21B, the human FKHR and AFX genes, identified as oncogene breakpoints but not as insulin signaling genes, are much more closely related to DAF-16 than the next closest relative in either Genbank or in the 94% complete *C. elegans* genome sequence. These data indicate that FKHR and AFX are excellent candidates for subserving the same function as *C. elegans* DAF-16: transduction of insulin signals and convergence with DAF-7-like Smad signals.

Evidence for the C. elegans AKT kinase as the probable output of DAF-2/AGE-1 signaling.

We screened genetically for mutations that bypass the need for age-1 signaling. This was done by mutagenizing a strain carrying an age-1(mg44)

null mutation (this mutation was heterozygous to allow the strain to grow). After two generations, animals that could survive without age-1 gene activity were selected by their lack of arrest at the dauer stage. We identified daf-16 mutations, as expected. However, we also identified two new gain of function mutations, sup(mg142) and sup(mg144).

sup(mg144) suppresses three different age-1 alleles, indicating that this mutation bypasses the need for AGE-1 production of PIP3. For example, sup(mg144) suppresses the dauer arrest of age-1(mg44), (m333), (mg109) such that fertile adults are formed. sup(mg144) does not suppress the lack of insulin signaling in the daf-2 mutant: daf-2(e1370); sup(mg144) form dauers at 25 degrees. This suggests that not all of the DAF-2 signaling output is via AGE-1. However, in the absence of both DAF-2 and AGE-1 signaling, sup(mg144) weakly suppresses, allowing some fertile adults to bypass arrest at the dauer stage. daf-2(e1370); sqt-1 age-1(mg44); sup(mg144) form 8% fertile adults, 12% sterile adults, and 80% dauers at 25 degrees.

Interestingly, sup(*mg144*) is a dominant suppressor of age-1 mutations. sqt-1 age-1(mg44); *sup(mg144)*/+ form 100% fertile adults. The sup(mg144) parental genotype does not affect this outcome. This data indicates that sup(*mg144*) is a dominant activating or dominant inactivating mutation.

Genetic mapping indicates that sup(*mg144*) may identify an activating mutation in the *C. elegans* AKT homologue (Fig. 25). By placing sup(mg144) in trans to a multiply marked chromosome (using PCR based RFLPs), we found that sup(mg144) maps to a 2 map unit genetic interval that includes *C. elegans* AKT (Fig. 24).

In particular, 2/39 sup(mg144) homozygous animals isolated from a sup(mg144)/polymorphic Bergerac chromosome parent recombined between sup(mg144)mg144 and stP6 (these animals also carried stP18). In this experiment mg144 was a heterozygote with RW7000 for three generations,

thus placing sup(mg144) approximately 2.2mu to the left of stP6.

In addition, 1/39 sup(mg144) homozygous animals isolated from a sup(mg144)/polymorphic Bergerac chromosome parent recombined between sup(mg144) and bP1. In this experiment mg144 was a heterozygote with RW7000 for two generations. Accordingly, this number is approximately 1/80 or 1.2 mu from bP1.

We generated a GFP fusion to AKT and showed that this gene is expressed at high levels in dauer larvae but at much lower levels and in fewer cells in wild type animals. (Figs. 26A-26B) Thus AKT represents a dauer regulated gene that may respond to DAF-16 and DAF-3 transcriptional control. Multiple probable binding sites, related to the DAF-3 binding site in myo11 have been identified.

sup(mg142) identifies another likely output of age-1 signaling

mg142 suppresses three different *age-1* alleles (*age-1(mg44)*, *age-1(m333)*, and *age-1(mg109)* at 20 degrees. *age-1(mg44)*; sup(mg142) form fertile adults at 15 and 20 degrees. At 25 degrees, they form 33% fertile adults and 67% sterile adults.

sqt-1 age-1(mg44); mg142/+ form 14% fertile adults and 86% sterile adults when the parent was homozygous for mg142. sqt-1 age-1(mg44); mg142/+ form 67% fertile adults and 33% sterile adults when the parent was heterozygous for mg142. daf-2(e1370); mg142 form sterile adults at 25 degrees; *daf-2(e1370)*; sqt-1 age-1(mg44); mg142 form sterile adults and dauers at 25 degrees. Preliminary mapping places mg142 approximately 1.6mu to the left of unc-1 on LGX.

Novel C. elegans insulin-like hormones are probable DAF-2 ligands

Mutations in *daf-2* not only cause a metabolic shift, but also affect

longevity of *C. elegans*. The nearly complete *C. elegans* genome sequence allowed a definitive search for insulin superfamily members to be performed, and, in this search, we detected multiple insulin-related proteins in the *C. elegans* genome database. When insulin, IGF-I, or IGF-II were compared to the translated worm genome sequence, this large set of insulin superfamily members was not detected. However, when the search was carried out with the conserved signature residues shown below that are the hallmark of the insulin superfamily (SEQ ID NOS: 115, 116), as now defined, we detected a number of novel insulin molecules.

Conserved Insulin Motifs

1 LCGXXLVEALXXVCGXRGFFYTPKTRRKRGIVEQCCXXXCXXXQL EXYCN 50 (SEQ ID NO: 115); and

1 aanqrLCGRHLADALYFVCGNRGFfyspkgGIVEECCHNPCTLYQLE NYCn 51 (an insulin superfamily consensus from the Blocks database at www.blocks.fhcrc.org; SEQ ID NO: 116).

The insulin superfamily signature residues were assembled using a set of vertebrate insulins and IGF-I and II proteins as well as silk moth bombyxin (a distant insulin relative) and a Limulus insulin superfamily member. The use of superfamily signature amino acid positions to detect distant relatives in databases is a more definitive approach to ascertaining gene superfamily members than simple searches with single family members.

Using these motifs, eight novel *C. elegans* insulin superfamily members were identified (SEQ ID NOS: 117-124), the coding sequences of which are shown in Figure 28. In this Figure, the family members are named from the cosmid genomic DNA sequences from which they were detected. All of these

insulins have A and B peptide homology to the insulin superfamily, and some of them have conserved dibasic processing sites that would mediate processing of the intervening unconserved C peptide. These genes are widely distributed on the *C. elegans* genome, although some are clustered (for example, ZK75.1, ZK75.2, ZK75.3, and ZK84.6). More distant insulin relatives may exist, but these are likely to engage receptors other than DAF-2.

Of the isolated insulin superfamily members, F13B12 was most closely related to human insulin and IGF-I, II. This was especially obvious from a PILEUP analysis in which a phylogenetic tree of protein superfamily members was constructed (Figs. 29 and 30). The insulin product of F13B12 clustered more closely to the mammalian insulin and IGF-I,II proteins than to other distant relatives like relaxin. Relaxin defined the most distantly related insulin superfamily member in the analysis, and it appeared to engage a tyrosine kinase receptor distinct from the insulin receptor.

These insulin-like hormones are expected to subserve the longevity, dauer arrest, and/or metabolic effects of DAF-2 signaling. For example, each of these insulin superfamily members are expected to engage the DAF-2 receptor, leading to a result in which a mutation in *daf-2* "sums" the functions of these eight or more insulin-like signals.

An analysis of the F13B12 insulin-like hormone is consistent with this view (Tables II-VI). First, as shown below, increasing the dose of the F13B12 insulin-like hormone potently modulates dauer arrest, both in animals carrying weak *daf-2* or weak *daf-7* mutations, and in animals carrying defects in synaptic components likely to mediate insulin release in *C. elegans* (unc-64).

Table II. High copy F13B12(ins) enhances the Daf-c phenotype of *daf*-2(*e*1365) at 20°C

Parental Genotype	Phenotype of progeny (%)					
	transgenic			non-transgenic		
	dauer	non-dauer	N	dauer	non-dauer	N
F13B12 transgenic: *daf*-2(*e*1365); *mgex*309	89.0	11.0	163	2.3	97.7	213
daf-2(*e*1365); *mgex*310	90.5	9.5	220	2.6	97.4	115
Control transgenic: *daf*-2(*e*1365); *mgex*315	1.8	98.2	283	0.5	99.5	184

Table III. High copy F13B12(ins) maternally suppresses the Daf-c phenotype of *daf*-7(*e*1372) at 25°C

Parental Genotype	Phenotype of progeny (%)					
	transgenic			non-transgenic (but parent was)		
	dauer	non-dauer	N	dauer	non-dauer	N
F13B12 transgenic: *daf*-7(*e*1372); *mgex*299	31.4	68.6	236	2.9	97.1	172
daf-7(*e*1372); *mgex*301	16.8	83.2	250	0	100	122
Control transgenic: *daf*-7(*e*1372); *mgex*312	100	0	78	100	0	60

Table IV. High copy F13B12(ins) maternally suppresses the Daf-c phenotype of *daf*-7(*e*1372) at 15°C

Parental Genotype	Phenotype of progeny (%)					
	transgenic			non-transgenic (but parent was)		
	dauer	non-dauer	N	dauer	non-dauer	N
F13B12 transgenic: *daf*-7(*e*1372); *mgex*299	1.4	98.6	73	0.3	99.7	343
daf-7(*e*1372); *mgex*301	0.5	99.5	194	0	100	278
Control transgenic: *daf*-7(*e*1372); *mgex*312	26.4	73.6	91	25.6	74.4	39

Table V. High copy F13B12(ins) promotes recovery of *unc*-64(*e*246) dauers at 27°C

Parental Genotype	Phenotype of progeny (%)						
	Day 2		Day 3				
			Transgenic		Non-transgenic		
	Dauer	Non-dauer	Dauer	Non-dauer	Dauer	Non-dauer	N
F13B12(ins) transgenic: *unc*-64(*e*246); *mgex*299	91.0	9.0	10.4	56.6	23.6	9.4	106
unc-64(*e*246); *mgex*301	75.3	24.7	22.9	51.1	18.7	7.3	96
Control transgenic: *unc*-64(*e*246); *mgex*312	88.9	11.1	54.3	10.6	29.8	5.3	208

Table VI. High copy F13B12(ins) enhances the Daf-c phenotype of *unc*-64(*e*246) at 15°C

Parental Genotype	Phenotype of progeny (%)					
	transgenic			non-transgenic		
	dauer	non-dauer	N	dauer	non-dauer	N
F13B12 transgenic: *unc*-64(*e*246); *mgex*299	23.2	76.8	185	0	100	170
unc-64(*e*246); *mgex*301	36.0	64.0	75	0	100	77
Control transgenic: *unc*-64(*e*246); *mgex*312	0	100	177	0	100	134

A genetic analysis has shown that high F13B12 insulin-like hormone signaling can suppress dauer arrest induced by *daf-7* mutations or decreases in synaptic signaling, but can enhance dauer arrest caused by decreases in *daf-2* signaling. Thus, the F13B12 insulin-like hormone may act synergistically with DAF-7 signals, like the DAF-2 receptor, but may interfere with the secretion or activity of another DAF-2 ligand. These genetic data strongly implicate the F13B12 insulin-like hormone in DAF-2 signaling.

In addition, the expression pattern of a promoter fusion of the F13B12 insulin-like hormone to GFP is also consistent with the genetic results. In these experiments, GFP was expressed in several head neurons, including ASJ and ASH, a pair of pharyngeal neurons, with processes that looked most like NSM, and three tail neurons. The full-length GFP looked similar but very faint. Worms expressing the full-length GFP lived longer than wild type. Interestingly, the NSM neuron had dense core vesicles by EM analysis, which is also true of beta cells of the pancreas. Pancreatic beta cells are also neuronal

in character; they use synaptic components for insulin vesicle release, are synaptically connected to the autonomic nervous system, and are electrically active. Sulfonyl ureas, which are used to increase insulin release, act by regulating the activity of K channels in beta cells, much the way K channels regulate excitability in other neurons. Finally, the NSM neuron is a part of the *C. elegans* enteric nervous system, just like the pancreas in mammals. Accordingly, the expression and functional analysis of the F13B12 insulin-like hormone is highly supportive of its role in insulin-like control of worm metabolism and aging.

Although the F13B12 insulin-like hormone is the closest *C. elegans* homologue to insulin, it is likely that many or all of these insulin superfamily members engage the DAF-2 receptor to regulate their activity. For example, they are more closely related to insulin than to the ligands of the other growth factor receptors present in the worm genome. These distinct insulin superfamily ligands could regulate DAF-2 at distinct times or places, or act antagonistically or synergistically to the F13B12 insulin-like hormone. Some of these insulin-like hormones may regulate metabolism, like insulin, whereas others may regulate dauer arrest or longevity. Thus, the *daf-2* mutant phenotype that results from loss of the receptor for these many hormones may be a composite loss of many hormonal signals. Consistent with such a model, neuronal expression of the DAF-2 receptor in a *daf-2* null mutant has been found to complement the dauer arrest phenotype of a *daf-2* mutant but not the metabolic or aging defects. Accordingly, one DAF-2 ligand may be expressed in or near the brain to control dauer arrest, but other ligands may impinge on DAF-2, for example, in non-neuronal cells, to control metabolism and aging.

By this view, loss of only one of the insulin-like hormones may cause only a subset of the *daf-2* mutant phenotype, for example, only increased longevity or only metabolic dysregulation. These *C. elegans* insulin

superfamily members may, for example, subserve the longevity or senescence function of DAF-2 receptor signaling, and an increase in such a hormone activity late in life may actually mediate the increase in DAF-2 activity that causes senescence. Conversely, if any of these insulin-like proteins have antagonistic effects on DAF-2, any decline in their activity late in life could mediate senescence. Application of only one hormone by injection or germ line therapy could therefore be used to target, for example, aging without any effects on metabolism.

In addition, since the F13B12 insulin-like hormone is a detectable worm homologue of insulin, it is possible that the other 7 worm insulins also have human homologues that are more closely related to their nematode counterparts than they are to each other. In fact the divergence of the F13B12 insulin-like hormone from insulin and IGF-I and IGF-II gives a measure of how much divergence may be expected for the mammalian homologues of the other insulin superfamily members. The F13B12 insulin-like hormone is slightly more closely related to IGF-II than insulin or IGF-I, but these three genes are probably duplicated and diverged homologues of a F13B12 homologue in the common ancestor of *C. elegans* and *Homo sapiens*. In fact, it is a current rule of thumb that many gene families in mammals have 4 times as many members as in *C. elegans*. For example, there are 4 Hox clusters in mammals and only one in *C. elegans*. Similarly, there are 3 known DAF-2 receptor homologues and DAF-16 transcription factor homologues in mammals (it is likely that the fourth mammalian member of these gene families will become known when the full mammalian genome sequence is finished). Thus, it is reasonable to expect that, for every insulin like protein in *C. elegans*, there may be four in mammals, or a total of 24 for the family of 8 shown above. In addition, since the F13B12 insulin-like hormone is expressed in only a few neurons, it is possible that the other insulin superfamily members are similarly expressed in a small set of

neurons, and that the human homologues may be expressed in only rare regulatory cell types.

The insulin-like hormones described herein, as well as their human homologues, provide valuable candidate regulators of senescence. For example, if human senescence is triggered by a decline in an insulin-like longevity hormone, in analogy to how puberty is triggered by a timed change in sexual maturation hormones, it may prove possible to regulate the aging process in the same way that sexual maturation can be regulated by hormone treatment. In addition, the *C. elegans* aging hormones may reveal which human genes have such a function. Because *daf-2* mutations cause longevity increases in a manner analogous to caloric restriction in mammals, it is possible that caloric restriction in mammals regulates the level of an insulin-like hormone that in turn engages the insulin or IGF-I, II receptors. Such a hormone may not have been detected if its level is very low or if it signals over a short range. However, once the human genome sequence is complete, the detection of human homologues to the *C. elegans* superfamily members listed above will become a trivial matter of database searching. In this way, the determination of the function of the worm homologue function in longevity or growth arrest or metabolism control will supply valuable functional information about the activity of human homologues.

The effect of the *C. elegans* insulin-like proteins on longevity, metabolism, or growth arrest may be readily determined by a combination of high copy studies, as shown above for the F13B12 insulin-like hormone, as well as by using RNA inhibition and knockout strategies to inhibit the activities of these genes. The *C. elegans* strains are then tested for interactions with *daf* pathway mutants, for example, as shown for the F13B12 insulin-like hormone above, and for longevity effects by standard techniques.

The human proteins that regulate longevity may be detected by a

combination of database searches and genetic complementation of worm RNAi or gene knockout mutants (for example, as described herein), as well as by high copy effects of human genes on worm longevity and metabolic control.

Because these human proteins are hormones, they may be used to directly regulate human longevity, for example, by injection into the bloodstream. Depending on the particular hormone and its effects, the hormones themselves may cause increased longevity, or they may be modified to generate dominant interfering hormones (for example, by engineering chimeras between the insulin superfamily members). The function of these proteins upon injection into the bloodstream may be predicted from their function in *C. elegans*, for example, as ascertained by transgenic analysis. Because of their effects on longevity, the human homologues of these *C. elegans* insulin-like endocrine signals have important applications in preventing or retarding the aging process.

C. elegans Akt/PKB Transduces Insulin Receptor-like Signals from AGE-1 Phosphoinositide-3-OH kinase to the DAF-16 Transcription Factor

An insulin receptor-like signaling pathway regulates *C. elegans* metabolism, development, and longevity (Kimura et al., *Science* 277:942-946, 1997). In response to a secreted pheromone, wild type animals arrest development at the dauer stage with a concomitant switch to fat storage metabolism in the intestine and hypodermis, increased lifespan, and remodelling of many tissues (Kimura et al., *Science* 277:942-946, 1997; Riddle and Albert, in C. elegans II, eds. Riddle, D.L., Blumenthal, T., Meyer, B.J. & Priess, J.R., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 739-768, 1997). Mutations in the insulin/IGF-I receptor homolog *daf-2* (Kimura et al., *Science* 277:942-946, 1997) or in the phosphoinositide-3-OH kinase (PI3K) homolog *age-1* (Morris et al., *Nature* 382:536-539, 1996) cause

constitutive arrest at the dauer stage; genetic analysis is consistent with AGE-1 functioning downstream of DAF-2 (Gottlieb and Ruvkun, *Genetics* 137:107-120, 1994; Larsen et al., *Genetics* 139:1567-1583, 1995). Mutations in the Fork head transcription factor DAF-16 completely suppress the dauer arrest, metabolic shift, and longevity phenotypes of *daf-2* and *age-1* mutants (Gottlieb and Ruvkun, *Genetics* 137:107-120, 1994; Larsen et al., *Genetics* 139:1567-1583, 1995; Kenyon et al., *Nature* 366:461-464, 1993; Ogg et al., *Nature* 389:994-999, 1997; Lin et al., *Science* 278:1319-1322, 1997), indicating that DAF-16 is a negatively regulated downstream target of *C. elegans* insulin receptor signaling. Molecules that couple the DAF-2 insulin receptor protein and AGE-1 PI3K to the DAF-16 transcription factor have not been identified by previous extensive genetic screens. While biochemical studies have suggested that the mammalian Akt/PKB (also known as RAC) serine/threonine kinase may transduce signals from PI3Ks associated with receptor tyrosine kinases (Franke et al., *Cell* 81:727-736, 1995; Burgering and Coffer, *Nature* 376:599-602, 1995; Cross et al., *Nature* 378:785-589, 1995), such as the insulin receptor to downstream effectors, this has not been demonstrated by genetic analysis of signaling pathways in whole organisms. We established the action of *C. elegans* Akt/PKB in the DAF-2 insulin receptor-like signaling pathway by the genetic identification of an activating Akt/PKB mutation and by genetic analysis of Akt/PKB inactivation and overexpression.

An activating mutation *(mg144)* in *akt-1*, one of two *C. elegans* Akt/PKB homologs, was identified in a genetic screen for mutations that suppress the dauer arrest phenotype of the *age-1(mg44)* null mutant (Morris et al., *Nature* 382:536-539, 1996). This screen was designed to isolate reduction of function mutations in molecules negatively regulated by PI3K signaling, or gain of function mutations in molecules positively regulated by PI3K signaling. Among 10 independent suppressor mutations isolated in a screen of 3800

haploid genomes, in addition to the activating *akt-1* mutation, we also isolated multiple alleles of a previously known negatively regulated target, *daf-16* (Gottlieb and Ruvkun, *Genetics* 137:107-120, 1994; Larsen et al., *Genetics* 139:1567-1583, 1995), suggesting that the screen revealed genes that act in this insulin-like signaling pathway.

The *mg144* mutation suppresses the three age-1 alleles tested, including two classes of nonsense alleles and one missense substitution (Ala845Thr) in a conserved region of PI3K (Morris et al., *Nature* 382:536-539, 1996). *mg144* is completely dominant for suppression of the dauer constitutive phenotype of *age-1(mg44)* (75.1% of the progeny of *age-1(mg44)*; *mg144*/+ animals developed as non-dauers, and 24.9% arrested at the dauer stage, N=774). On its own, *mg144* does not have any obvious phentoypes; it moves normally, has a normal vulva and brood size, and makes dauers on starved plates and on plates treated with pheromone. Thus *mg144* does not activate the AGE-1 PI3K signaling pathway to the point that normal dauer arrest is affected but does activate the pathway sufficiently to alleviate the requirement for AGE-1 PI3K outputs.

Using suppression of the dauer constitutive phenotype of *age-1(mg44)*, *mg144* was mapped to a region on chromosome V within 1.3mu of the polymorphic STS marker bP1 (Fig. 31). From the *C. elegans* genome sequence in this 1.3 mu region, we identified a *C. elegans* Akt/PKB homolog which we named akt-1 (Fig. 31). Because an activating mutation in Akt/PKB is a good candidate to be a genetically dominant suppressor of an age-1 PI3K null mutant, we determined the *akt-1* DNA sequence in the *mg144* strain by PCR amplification and direct sequencing. The *akt-1* gene in the *mg144* mutant strain was shown to bear an Ala183Thr substitution (Fig. 34). *akt-1* is differentially spliced within the conserved kinase domain to generate the *akt-1a* and *akt-1b* isoforms with distinct kinase domain subregions IV, V, and VI (13)

(92% identical, 238/258 amino acids over the entire kinase domain; 69% identical, 44/64 amino acids in the differentially spliced region). *akt-1a* is 58% identical to human Akt/PKBa (Fig. 33 and 34). *akt-1* has a pleckstrin homology domain, kinase domain, and the two phosphorylation sites necessary for Akt/PKB activation (Alessi et al., *EMBO J.* 15:6541-6551, 1996) which are the hallmarks of the Akt/PKB family (Fig. 34). The next most closely related non-Akt/PKB mammalian kinase is rat PKCb1 which is 38% identical to *akt-1a*. The *akt-1*(*mg144*) mutation is present in both splice forms of *akt-1* and is located in a region of the protein that links the N-terminal pleckstrin homology domain to the C-terminal kinase domain. This mutation is in a region that is not conserved between *C. elegans* and mammalian Akt/PKB. This mutation may reveal a negative regulatory region on *akt-1* because the *mg144* allele is an activating mutation (see below).

To confirm that the *mg144* suppression of *age-1* that is genetically linked to *akt-1* was due to a mutation in *akt-1*, we used a reverse genetic assay termed RNA interference (RNAi) (Fire et al., *Nature* 391:806-811, 1998; Rocheleau et al., *Cell* 90:707-716, 1997; Zhang et al., *Nature* 390:477-484, 1997) to decrease *akt-1* gene activity in an *age-1*(*mg44*); *akt-1*(*mg144*) strain. If a mutation in *akt-1* was responsible for the suppression of *age-1* observed in this strain, RNAi of *akt-1* in this strain should revert the suppression phenotype and result in a dauer constitutive phenotype. This experiment was conceptually similar to the classic genetic arguments that show that a cis-acting loss of function mutation can revert a gain of function mutation in the same gene. Inhibition of *akt-1* activity in an *age-1*(*mg44*); *akt-1*(*mg144*) strain reverted the *akt-1*(*mg144*) suppression phenotype, indicating that the *mg144* activating mutation was a lesion in the *akt-1* locus.

We identified another Akt/PKB homolog in the nearly complete *C. elegans* genome sequence (Wilson et al., *Nature* 368:32-38, 1994) which we

named *akt-2* (Fig. 32). *akt-1* and *akt-2* are more closely related to each other (66% identity between *akt-1a* and *akt-2* overall) than to any other Akt/PKB homolog (Fig. 33). *akt-2* is 55% identical to human Akt/PKBa overall and 35% identical to rat PKCb1 overall. Interestingly, *akt-2* only has the Thr308 phosphorylation site that is necessary for Akt/PKB activation by PDK1 (Alessi et al., *Current Biology* 7:261-269, 1997; Stokoe et al., *Science* 277:567-570, 1997) but not the Ser473 phosphorlyation site (Alessi et al., *EMBO J.* 15:6541-6551, 1996) (Fig. 34) and yet clearly functions in the insulin-like signaling pathway (see below).

Reduction of both *akt-1* and *akt-2* activities revealed that they transduce insulin-like signals from the AGE-1 PI3K to the DAF-16 forkhead transcription factor. Inhibition of either *akt-1* or *akt-2* activity by RNAi did not cause dauer arrest. However, simultaneous inhibition of both *akt-1* and *akt-2* activities caused nearly 100% arrest at the dauer stage. We concluded that Akt/PKB signaling from either *akt-1* or *akt-2* is sufficient for reproductive development. This result indicates that *akt-1* and *akt-2* can function redundantly for dauer formation in *C. elegans* and raises the possibility that various mammalian Akt/PKB isoforms could function redundantly as well. Significantly, the constitutive dauer arrest induced by inhibition of both *akt-1* and *akt-2* is fully suppressed by a null mutation in *daf-16* (Ogg et al., *Nature* 389:994-999, 1997) but is not suppressed by a null mutation in the Smad homolog *daf-3* (Patterson et al., *Genes & Development* 11:2679-2690, 1997) which confirms its placement in the DAF-2/AGE-1/DAF-16 signaling pathway. Because a null mutation in *daf-16* alleviates the need for *C. elegans* Akt/PKB signaling, the primary function of AKT-1 and AKT-2 is to antagonize DAF-16. Interestingly, DAF-16 contains four consensus sites for phosphorylation by Akt/PKB (Alessi et al., *FEBS Letters* 399:333-338, 1996) and three of these sites are conserved in the human DAF-16 homologs AFX, FKHR, and FKHRL1. AKT-1 and

AKT-2 may exert their negative regulatory effect by directly phosphorylating DAF-16. Shown below are comparisons of AFX, FKHR, and DAF-16, indicating the conservation between the consensus phosphorylation sites. The AKT sites indicated are located downstream and upstream, respectively, of the Forkhead domain SEQ ID NOS: 161-169).

Score = 151 (68.4 bits), Expect = 1.9e-140, Sum P(8) = 1.9e-140
Identities = 28/54 (51%), Positives = 38/54 (70%)

```
AFX:    226 SPVGHFAKWSGSPCSRNREEADMWTTFRPRSSSNASSVSTRLSPLRPESEVLAE 279
            SP   F+KW  SP S + ++ D W+TFRPR+SSNAS++S RLSP+  E + L E
FKHR:   287 SPGSQFSKWPASPGSHSNDDFDNWSTFRPRTSSNASTISGRLSPIMTEQDDLGE 340
DAF-16a                                SFRPRTQSNLSIPGSSS
```

Score = 132 (59.8 bits), Expect = 1.9e-140, Sum P(8) = 1.9e-140
Identities = 22/42 (52%), Positives = 28/42 (66%)

```
AFX:      7 KAAAIIDLDPDFEPQSRPRSCTWPLPRPEIANQPSEPPEVEP 48
            +A  ++++DPDFEP  RPRSCTWPLPRPE +   S       P
FKHR:     3 EAPQVVEIDPDFEPLPRPRSCTWPLPRPEFSQSNSATSSPAP 44
DAF-16  TFMNTPDDVMMNDDMEPIPRDRCNTWPMRRPQLEPPLNSSP 177
        T  ++P+ V ++ D EP+PR R  TWP+ RP++  + ++++
```

We have shown that human AKT will phosphorylate *C. elegans* DAF-16 and that this phosphorylation is dependent on these sites. Upon mutation of the serine or threonine in these sites to alanine, *in vitro* phosphorylation of DAF-16 (or fragments of DAF-16) is abolished. It is expected that the lack of *akt* input to DAF-16 in these mutant nematodes will result in dauer arrest, just like animals lacking *akt-1/akt-2* gene activity.

The above genetic results show that Akt/PKB is the major output of PI3K signaling and implicate a transcription factor downstream target for the Akt/PKB kinase. Because mutations in *daf-16* suppress *akt-1* and *akt-2* reduction of function, it is likely that DAF-16 represents a major signaling

output of Akt/PKB in *C. elegans* insulin-like signaling. Akt/PKB has been implicated in mammalian insulin receptor signaling that localizes glucose transporters to the plasma membrane (Kohn et al., *J. Biol. Chem.* 271:31372-31378, 1996) and has been shown to regulate glycogen synthesis via direct phosphorylation of GSK-3 (Cross et al., *Nature* 378:785-589, 1995), two events which are not transcriptionally regulated. While there also may be such Akt/PKB outputs in *C. elegans*, the DAF-16 Fork head transcription factor represents the major output of DAF-2/AGE-1/AKT-1/AKT-2 insulin receptor-like signaling (Ogg et al., *Nature* 389:994-999, 1997). Similarly Akt/PKB action in the insulin/IGF-I anti-apoptotic pathway (Dudek et al., *Science* 275:661-665, 1997; Kauffmann-Zeh et al., *Nature* 385:544-548, 1997; Kulik et al., *Mol. Cell Biol.* 17:1595-1606, 1997 24-26) may also converge on transcription factors related to DAF-16.

The normal requirement of *age-1* activity for reproductive development is also bypassed by increased gene dosage of wild type *akt-1*. Transgenic *age-1*(*mg44*) animals carrying a 7.3 kb *akt-1*(+) genomic region can grow reproductively rather than arrest at the dauer stage. Greater than 75% of *age-1*(*mg44*) animals that contain the *akt-1*(+) transgene at high copy bypass dauer arrest while non-transgenic *age-1*(*mg44*) animals never bypass dauer arrest. In a similar experiment with *age-1*(*mg44*) animals carrying the same genomic region amplified from *akt-1*(*mg144*) at high copy, the transgenic animals bypassed dauer arrest at a similar frequency. The *age-1*(*mg44*) animals carrying the *akt-1*(*mg144*) transgene at low copy bypass dauer arrest more frequently than the *age-1*(*mg44*) animals carrying the *akt-1*(+) transgene at low copy (approximately 85% of *age-1*(*mg44*) animals carrying *akt-1*(*mg144*) transgene bypass dauer compared to 38% of *age-1*(*mg44*) animals carrying the *akt-1*(+) transgene). These results indicate that the same 7.3 kb genomic region amplified from the *akt-1*(*mg144*) strain is a more potent suppressor of

age-1(*mg44*) than the *akt-1*(+) transgene. These data map *mg144* to the 7.3 kb region of akt-1 that includes the Ala183Thr substitution in AKT-1. These data also suggest that the mutation may act by increasing AKT-1 abundance or stability, thus conferring the ability to grow in the absence of age-1 signaling.

Null mutations in age-1 cause dauer arrest as does inactivation of *akt-1* and *akt-2* by RNAi. This indicates that *akt-1*(+), *akt-2*(+), and *age-1*(+) are required for reproductive development. Because the dominant allele *akt-1*(*mg144*) also promotes reproductive growth by virtue of its ability to suppress the dauer constitutive phenotype of *age-1* null mutants, it functions similarly to *akt-1*(+) and *akt-2*(+). Thus *akt-1*(*mg144*) is an activating mutation, as opposed to a loss of function or dominant negative mutation in *akt-1*. In addition, the fact that both *akt-1*(*mg144*) and providing additional copies of the *akt-1*(+) gene suppress an *age-1* null mutant is consistent with *akt-1*(*mg144*) being an activating mutation.

akt-1(*mg144*) suppresses the dauer constitutive phenotype of three *age-1* alleles. Because *age-1*(*mg44*) is a null mutant, these data strongly suggest that *akt-1* acts downstream of *age-1* and demonstrates that the biochemical ordering of PI3K upstream of Akt/PKB kinase is also true in an intact organism. AGE-1 is the only PI3K homolog in *C. elegans* of the type regulated by tyrosine kinase receptors. Significantly, our results demonstrate that *C. elegans* Akt/PKB gene activity is not strictly dependent on upstream *age-1* activity if Akt/PKB activity is increased because *akt-1*(*mg144*) as well as *akt-1*(+) overexpression suppress null mutations in AGE-1 PI3K. This is comparable to the suppression by *daf-16(m27)*, a reduction of function allele (Lin et al., *Science* 278:1319-1322, 1997), and *daf-16* null alleles (Ogg et al., *Nature* 389:994-999, 1997).

A mutation in *daf-2* is suppressed more poorly by *akt-1*(*mg144*) than by a reduction of function mutation in *daf-16*. The *age-1* alleles suppressed by *akt-1*(*mg144*) are null (Morris et al., *Nature* 382:536-539, 1996) whereas

daf-2(e1370) is a temperature sensitive mutation in the kinase domain (Kimura et al., *Science* 277:942-946, 1997). This *daf-2* allele is completely suppressed by many *daf-16* alleles, including null alleles (Gottlieb and Ruvkun, *Genetics* 137:107-120, 1994; Larsen et al., *Genetics* 139:1567-1583, 1995; Ogg et al., *Nature* 389:994-999, 1997). Because *akt-1(mg144)* can bypass the need for AGE-1 PI3K signaling but not for DAF-2 insulin receptor-like signaling, *akt-1(mg144)* defines a bifurcation in the signaling pathway downstream of *daf-2*. It is likely that *age-1* and *akt-1* constitute one major signaling pathway from DAF-2 and that other, as yet unidentified genes, constitute one or more parallel pathways. These pathways converge downstream of AGE-1 and at or upstream of the DAF-16 Fork head transcription factor and negatively regulate its activity, since loss of function mutations in *daf-16* completely suppress both *daf-2* and *age-1* mutations (Gottlieb and Ruvkun, *Genetics* 137:107-120, 1994). Because a decline in AGE-1 PI3K or AKT-1/AKT-2 signaling induces dauer arrest in the presence of signaling from this parallel pathway, both are necessary for reproductive development. The genetic evidence for multiple DAF-2 insulin receptor-like outputs demonstrate that biochemical studies showing that parallel PI3K, ras, SHP2, and other signaling outputs are activated by the insulin receptor in mammals (Kahn, *Diabetes* 43:1066-1084, 1994) are relevant to insulin receptor-like signaling in intact organisms.

Reduction of zygotic *age-1* activity increases *C. elegans* lifespan greater than two-fold (Morris et al., *Nature* 382:536-539, 1996; Larsen et al., *Genetics* 139:1567-1583, 1995; Klass, *Mech. Ageing Dev.* 22:279-286, 1983). Mutations in *daf-16* suppress this lifespan increase (Larsen et al., *Genetics* 139:1567-1583, 1995; Dorman et al., *Genetics* 141:1399-1406, 1995). *akt-1(mg144)* does not suppress the *age-1(mg44)* induced increase in lifespan (for the following strains, mean lifespans, maximum lifespan are given: N2 12 days, 16 days, N=28; *sqt-1(sc13) age-1(mg44)* 18 days, 36 days, N=20;

sqt-1(sc13) age-1(mg44); *akt-1(mg144)* 22 days, 38 days, N=36; *daf-16(m27)*; *sqt-1 (sc13) age-1(mg44)* 14 days, 16 days, N=32). Thus *akt-1*(*mg144*) bypasses the need for AGE-1 signaling in reproductive development but does not activate normal aging pathways. It is possible that *akt-1*(*mg144*) does not subserve all the functions of the wild type *akt-1* or *akt-2*. *akt-2* or other as yet unidentified downstream effectors of *age-1* may be the pertinent signaling molecules for lifespan regulation.

The expression patterns of both *akt-1* and *akt-2* were examined in transgenic animals containing a translational fusion of each genomic locus to Green Fluorescent Protein (GFP) (Chalfie et al., *Science* 263:802-805, 1994). The GFP fusion proteins contain the entire genomic coding region from either *akt-1* or *akt-2*, including 5' upstream regulatory sequence, fused in frame at the C-terminus to GFP. AKT-1/GFP expression is first observed in late embryos and is maintained throughout the life of the animal. In post-embryonic animals, AKT-1/GFP is expressed in the majority of head neurons including sensory neurons. Expression is also observed in motor neurons of the ventral and dorsal nerve cord, neuronal commissures and processes throughout the body, and the tail neurons. The fusion protein is localized throughout the cell body and axonal and dendritic processes of neurons but is usually excluded from the nucleus. Additional tissues which consistently express AKT-1/GFP include neurons and muscle cells of the pharynx, the rectal gland cells, and the spermatheca. AKT-1/GFP expression was observed more variably in a variety of cell types including hypodermis, intestine, muscle, some of the P cell descendants that form the vulva, and in a structure we believe to be the excretory canal. Consistent with redundant roles of *akt-1* and *akt-2*, an AKT-2/GFP full length protein fusion gene is expressed at the same times as AKT-1/GFP and in the same tissues that express AKT-1/GFP, although AKT-2/GFP seems to be less abundant. In dauers induced by starvation on

crowded plates, AKT-1/GFP and AKT-2/GFP expression does not differ dramatically from their expression during reproductive growth. These expression patterns are consistent with AKT-1 and AKT-2 functioning either in secretory neurons to regulate dauer arrest and metabolic shift or in the target tissues that are remodeled during dauer formation.

The role of AKT-1 and AKT-2 in regulating the metabolic shift and developmental arrest associated with dauer formation suggests the following model. Under normal growth conditions, an insulin-like molecule binds to the DAF-2 insulin receptor kinase inducing autophosphorylation and recruitment of AGE-1 PI3K. As discussed herein, PI3K signals via Akt/PKB. Precedent from biochemical experiments in other systems (Franke et al., *Cell* 81:727-736, 1995; Franke et al., *Science* 275:665-668, 1997; Klippel et al., *Mol. Cell Biol.* 17:338-344, 1997) suggests that AGE-1 activation produces phospholipids that bind to and activate AKT-1 and AKT-2 by inducing a conformational change in the protein that makes it accessible to phosphorylation events which are necessary for activation (Alessi et al., *Current Biology* 7:261-269, 1997; Stokoe et al., *Science* 277:567-570, 1997). A parallel pathway or pathways from the DAF-2 insulin receptor-like protein is also activated. The AKT-1 and AKT-2 kinases, as well as molecules from the parallel pathway, negatively regulate DAF-16 activity, possibly via phosphorylation. Phosphorylated DAF-16 could be inactive, function to activate genes required for reproductive growth and metabolism, or repress genes required for dauer arrest and energy storage. Other signaling molecules that are activated by DAF-2 must also converge downstream of AGE-1 (for example, on DAF-16 or AKT-1/AKT-2) for proper regulation of metabolism and lifespan: the dauer arrest induced by loss of AGE-1 PI3K or AKT-1/AKT-1 activity implies that the loss of only one of these inputs to DAF-16 is sufficient to cause dauer arrest. Under dauer inducing conditions, DAF-2, AGE-1, AKT-1/AKT-2, and other signaling

pathways from DAF-2 are inactive and therefore DAF-16 is active, presumably because it is under-phosphorylated. Active DAF-16 either represses genes required for reproductive growth and metabolism or activates genes necessary for dauer arrest and energy storage.

The DAF-16 Fork head protein has been suggested to interact with the DAF-3, DAF-8, or DAF-14 Smad proteins to integrate converging TGF-β like neuroendocrine signals with insulin-like signals (Ogg et al., *Nature* 389:994-999, 1997; Patterson et al., *Genes & Development* 11:2679-2690, 1997). DAF-16 may form a complex with the DAF-3 Smad protein under dauer inducing conditions to regulate these downstream genes (Ogg et al., *Nature* 389:994-999, 1997), while AKT-1 phosphorylation of DAF-16 may inhibit the formation of a Smad/Fork head complex during reproductive development.

This model, based on genetic evidence that Akt/PKB couples insulin receptor-like signaling to transcriptional output via the DAF-16 Fork head transcription factor in *C. elegans*, predicts that Akt/PKB will have transcriptional outputs in insulin-like signaling across phylogeny. It was previously suggested that the human homologs of the DAF-16 transcription factor (AFX, FKHR, FKHRL1 and AF6q21) may be the pertinent downsteam effectors of insulin signaling in humans (Ogg et al., *Nature* 389:994-999, 1997). A recent report shows that Akt/PKB mediates insulin dependent repression of the insulin-like growth factor binding protein-1 (IGFBP-1) gene in HepG2 cells via a conserved insulin response sequence (CAAAAC/TAA) (Cichy et al., *J. Biol. Chem.* 273:6482-6487, 1998). Interestingly, we have determined that DAF-16 binds to this same insulin response sequence *in vitro*. We propose that Akt/PKB mediates its transcriptional effects on insulin responsive genes such as IGFBP-1 via the human homologs of DAF-16: AFX, FKHR, FKHRL1, or AF6q21.

PDK genetics

From the same genetic screen that generated the *akt-1(mg144gf)* allele, we identified another *age-1* suppressor, *mg142*. This mutation also bypasses the need for upstream *age-1* signaling and is genetically dominant. Genetic mapping placed the mutation in the region where a *C. elegans* homologue maps. The genomic sequence of *pdk-1*, starting 60 bp upstream of the start codon and ending 60 bp downstream of the stop codon is shown in Figure 35 (SEQ ID NO: 158). Figures 36 and 37 show the two *C. elegans pdk-1* spliced forms, *pdk-1a* (Figure 36; SEQ ID NO: 159) and *pdk-1b* (Figure 37; SEQ ID NO: 160). The *pdk-1(mg142)* gain of function mutation is Ala303Val (splice 1). This protein is 58% identical to mammalian PDK in the pleckstrin homology domain and 39% identical in the kinase domain as shown below SEQ ID NOS: 170-199).

```
 Score = 252 (88.7 bits), Expect = 2.2e-60, Sum P(6) = 2.2e-60
 Identities = 47/80 (58%), Positives = 60/80 (75%), Frame = +3

Query:   439 LEKQAGGNPWHQFVENNLILKMGPVDKRKGLFARRRQLLLTEGPHLYYVDPVNKVLKGEI 498
             LE+Q   NP+H F  N+LILK G ++K++GLFARRR  LLTEGPHL Y+D  N VLKGE+
Sbjct:  1818 LEEQRVKNPFHIFTNNSLILKQGYLEKKRGLFARRRMFLLTEGPHLLYIDVPNLVLKGEV 1997

Query:   499 PWSQELRPEAKNFKTFFVHT 518
             PW+  ++ E KN  TFF+HT
Sbjct:  1998 PWTPCMQVELKNSGTFFIHT 2057

 Score = 201 (70.8 bits), Expect = 2.2e-60, Sum P(6) = 2.2e-60
 Identities = 48/123 (39%), Positives = 72/123 (58%), Frame = +1

Query:   263 SDLWALGCIIYQLVAGLPPFRAGNEYLIFQKIIKLEYDFPEKFFPKARDLVEKLLVLDAT 322
             +D+W LGCI++Q +AG PPFRA N+Y + ++I +L++ FPE F  +A +++ K+LV
Sbjct:   802 TDIWGLGCILFQCLAGQPPFRAVNQYHLLKRIQELDFSFPEGFPEEASEIIAKILV-G*H 978

Query:   323 KRLGCE----EMEGYGP--------LKAHPFFESVTWENLHQQTPPKLTAYLPAMSEDDE 370
             + L  E     ++   P         L AH FFE+V W N+    PP L AY+PA   + E
Sbjct:   979 ETLKTEYVIFNLQVRDPSTRITSQELMAHKFFENVDWVNIANIKPPVLHAYIPATFGEPE 1158

Query:   371 DCYGN 375
               Y N
Sbjct:  1159 -YYSN 1170

 Score = 180 (63.4 bits), Expect = 2.2e-60, Sum P(6) = 2.2e-60
 Identities = 31/72 (43%), Positives = 52/72 (72%), Frame = +2

Query:   157 FGLSYAKNGELLKYIRKIGSFDETCTRFYTAEIVSALEYLHGKGIIHRDLKPENILLNED 216
             F +   +NG+L + +   GSFD   ++F+ +EI++ L++LH   I+HRD+KP+N+L+ +D
Sbjct:   287 FVIGLVENGDLGESLCHFGSFDMLTSKFFASEILTGLQFLHDNKIVHRDMKPDNVLIQKD 466
```

```
Query:    217 MHIQITDFGTAK 228
              HI ITDFG+A+
Sbjct:    467 GHILITDFGSAQ 502

 Score = 83 (29.2 bits), Expect = 2.2e-60, Sum P(6) = 2.2e-60
 Identities = 15/53 (28%), Positives = 32/53 (60%), Frame = +2

Query:    108 YAIKILEKRHIIKENKVPYVTRERDVMSRLD-----HPFFVKLYFTFQDDEKL 155
              +A+K+L+K ++ +  K+  + RE+++++ L      HPF  +LY  F D  ++
Sbjct:      8 FAVKVLQKSYLNRHQKMDAIIREKNILTYLSQECGGHPFVTQLYTHFHDQARI 166

 Score = 81 (28.5 bits), Expect = 2.2e-60, Sum P(6) = 2.2e-60
 Identities = 15/29 (51%), Positives = 19/29 (65%), Frame = +2


Query:    519 PNRTYYLMDPSGNAHKWCRKIQEVWRQRY 547
              PNR YYL D     A +WC+ I +V R+RY
Sbjct:   2129 PNRVYYLFDLEKKADEWCKAINDV-RKRY 2212

 Score = 78 (27.5 bits), Expect = 2.2e-60, Sum P(6) = 2.2e-60
 Identities = 15/25 (60%), Positives = 18/25 (72%), Frame = +3

Query:    232 PESKQARANSFVGTAQYVSPELLTE 256
              PE   AR  +FVGTA YVSPE+L +
Sbjct:    660 PEENTARRTTFVGTALYVSPEMLAD 734
```

Mapping of the *mg142* mutation to this open reading frame establishes the function of this protein. It is much more closely related to PDK than to any other known kinase. PDK is a mammalian kinase that phosphorylates an essential serine residue on AKT, contributing to its activation. This serine is conserved in *akt-1* and *akt-2*. Thus, PDK is an excellent candidate gene for the *mg142* mutation. The genetic region bearing *pdk-1* was amplified from the *mg142* strain, and an amino acid substitution in a conserved region of the PDK kinase domain was detected. While a gain of function mutation in *pdk* would be consistent with the biochemical work that shows that PDK acts upstream of AKT to activate it, this genetic work suggests that, if PDK can be activated (for example, by the *mg142* mutation), no PIP3 signaling from the AGE-1 PI3K is necessary, since *mg142* suppresses an *age-1* null allele. To establish that this substitution causes the suppression of *age-1* induced dauer arrest, a strategy analogous to that used to analyze the *akt-1*(*mg144*gf) mutation may be utilized.

Because we have implicated PDK in the *C. elegans* insulin signaling pathway, human PDK1 becomes a candidate gene for variation in diabetes.

Mutations in human PDK1 may underlie the genetic variation that causes diabetes in some families. Similarly, drugs that activate PDK, like the *mg142* mutation that activates *C. elegans pdk-1*, may bypass the need for upstream signaling in some diabetics with such upstream defects. The region of human PDK1 that is homologous to the *C. elegans pdk-1* at alanine 303 provides a good candidate for screening for drugs that bind and activate signaling. Similarly, the region of human AKT between the kinase domain and the PH domain, where the *C. elegans akt-1* gain of function mutation maps is a good candidate for the design of drugs that activate AKT. Such activated AKT in *C. elegans* bypasses the need for upstream signaling from the AGE-1 PI3K and may similarly treat diabetics with defects in insulin signaling between insulin and AKT.

Diapause and Longevity

Weak *daf-2* and *age-1* mutants that do not arrest at the dauer stage nevertheless live much longer than wild-type (Larsen et al., *Genetics* 139: 1567-1583, 1995; Kenyon et al., *Nature* 366: 461-464, 1993; Dorman et al., *Genetics* 141: 1399-1406, 1995). This connection between longevity and diapause control may not be unique to *C. elegans*. Diapause arrest is an essential feature of many vertebrate and invertebrate life cycles, especially in regions with seasonal temperature and humidity extremes (Tauber et al., *Seasonal Adaptation of Insects*, Oxford University Press, New York, N. Y., 1986). Animals in diapause arrest slow their metabolism and their rates of aging, and can survive for periods for much longer than their reproductive lifespan (Tauber et al., *supra*, 1986).

Because insulin-like DAF-2/AGE-1 signaling mediates *C. elegans* diapause longevity control, the mammalian insulin signaling pathway may also control longevity homologously. In fact, the increase in longevity associated

with decreased DAF-2 signaling is analogous to mammalian longevity increases associated with caloric restriction (Finch, *Longevity, Senescence and the Genome*, The University of Chicago Press, Chicago, 1990). It is possible that caloric restriction causes a decline in insulin signaling to induce a partial diapause state, like that induced in weak *daf-2* and *age-1* mutants. The induction of diapause-like states may affect post-reproductive longevity (Finch, *supra*), as in *C. elegans*. Alternatively, it is the changes in the mode and tempo of metabolism itself rather than diapause per se that causes increased longevity. Another long-lived *C. elegans* mutant, *clk-1*, may also regulate lifespan via such metabolic effects (Ewbank et al., *Science* 275: 980-983, 1997). This association of metabolic rate with longevity is also consistent with the correlation of free radical generation to aging (Finch, *supra*).

Synergistic Control of Metabolism and Diapause by Insulin and TGF-β Signaling Pathways

In addition to DAF-2 signaling, the DAF-7 TGF-β neuroendocrine signal is also necessary for reproductive development of *C. elegans* (Ren et al., *Science* 274: 1389-1391, 1996; Schackwitz et al., *Neuron* 17: 719-728, 1996). The signals in these two pathways are not redundant: animals missing either *daf-2* signaling or *daf-7* signaling (Fig. 3) shift their metabolism and arrest at the dauer stage (Table VII). In addition the phenotypes caused by mutations in either pathway are strongly synergistic, suggesting that the two pathways are integrated. Synchronised eggs were grown and counted as described above. *daf-1(m40)* and *daf-2(e1370)* form 100% dauer at 25°C. Numbers shown in Table VII indicate percentage dauer formation and number of animals counted (in parenthesis). Data presented is the sum of three independent trials.

Table VII. Synergy of *daf-1* and *daf-2*

	% dauer formation	
	15°C	20°C
daf-1 (m40)	0.0 (532)	1.9 (909)
daf-2 (e1370)	0.0 (798)	3.8 (503)
daf-1 (m40); *daf-2* (e1370)	19.4 (747)	100 (718)

This data indicates that DAF-7 TGF-β signals and DAF-2 ligand insulin-like signals are integrated. In support of this model, weak mutations in the *daf-2* insulin signaling pathway and in the *daf-7* TGF-β signaling pathway are highly synergistic (Table VII). Genetic epistasis analysis indicates that the DAF-7 and DAF-2 pathways are parallel rather than sequential (Vowels and Thomas, *Genetics* 130: 105-123, 1992; Gottlieb and Ruvkun, *Genetics* 137: 107-120, 1994). That is, *daf-16* mutations strongly suppress *daf-2* mutations but not *daf-7*, *daf-1*, or *daf-4* mutations, whereas *daf-3* mutations strongly suppress *daf-7*, *daf-1*, and *daf-4* mutations, but not *daf-2* mutations. Analogous synergism between activin and FGF tyrosine kinase pathways in *Xenopus* mesoderm induction has been noted (Green et al., *Cell* 71: 731-739, 1992).

A dauer-inducing pheromone regulates the production of DAF-7 by the ASI sensory neuron (Ren et al., *Science* 274: 1389-1391, 1996; Schackwitz et al., *Neuron* 17: 719-728, 1996). Because animals carrying *daf-7* nonsense or truncation mutations are responsive to pheromone (Golden and Riddle, *Proc. Natl. Acad. Sci. U.S.A. 81*: 819-823, 1984), we further suggest that the production of the insulin-like ligand for DAF-2 is also regulated by pheromone. It is not yet clear whether these DAF-7 and DAF-2 signals converge in target tissues or in other regulatory (i.e., hormonal) cells; however the expression of

the DAF-7 receptor pathway genes in essentially all target tissues (infra) suggests that integration occurs there.

DAF-7 and Diabetes

Based on the data herein, we propose that in humans as in *C. elegans*, both a DAF-7-like neuroendocrine signal and insulin are necessary for metabolic control by insulin. According to this model, the failure of target tissues to respond to insulin signals in Type II diabetic patients could be due to defects either in the insulin or TGF-β-like control pathways. Pedigree analysis has shown a strong genetic component in Type II diabetes (Kahn et al., *Annu. Rev. Med.* 47: 509-531,1996). In addition, obesity is also a major risk factor in Type II diabetes (Kahn et al., *Annu. Rev. Med.* 47: 509-531,1996). Genetic or obesity-induced (Hotamisligil et al., *Science* 259: 87-91, 1993; Lonnqvist et al., *Nat Med* 1: 950-953, 1995) declines in a DAF-7-like signaling pathway could underlie the lack of response to insulin in Type II diabetes, just as in *C. elegans daf-7* mutants cause metabolic defects very similar to *daf-2* mutants. The discovery that the DAF-7 and DAF-2 pathways converge indicates that DAF-7 hormonal signals are defective in diabetic conditions (for example, Type II diabetes), and that administration of human DAF-7 is useful for ameliorating the glucose intolerance, ketoacidosis, and atherosclerosis associated with diabetes. This is shown schematically in Figs. 17, 18, and 23.

Whereas the DAF-7 TGF-β like and DAF-2 insulin-like signaling pathways converge to control diapause and metabolism, only the DAF-2/AGE-1 pathway has been implicated in reproductive adult stage longevity control in the absense of dauer formation (Larsen et al., *Genetics* 139: 1567-1583, 1995; Kenyon et al., *Nature* 366: 461-464, 1993; Dorman et al., *Genetics* 141: 1399-1406, 1995; and Morris et al., *Nature* 382: 536-539, 1996). Both pathways control the longevity increase associated with dauer arrest, since

dauer larvae live much longer than reproductive *C. elegans* (Riddle, In: *Caenorhabditis elegans* II, D. Riddle, T. Blumenthal, B. Meyer, J. Priess, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1997, pp. 739-768; Kenyon, op cit. pp., 791-813: Chayen and Bitensky, *Practical Histochemistry*, Chichester; New York: Wiley, 1991. The distinction between DAF-7 and DAF-2 regulation of longevity could also reflect a more profound regulation of metabolism by the DAF-2 pathway than the DAF-7 pathway (Fig. 4). For example, based on precedents from TGF-β signaling in other systems and analysis of this pathway in *C. elegans*, all of the known signaling output of the DAF-7 TGF-β pathway are via downstream Smad transcriptional regulation (infra). Insulin signaling, and by extension, DAF-2 signaling, is more ramified: outputs from this receptor regulate sugar transport, metabolic enzyme activities, translation of mRNAs encoding these and other enzymes, as well as transcription (White and Kahn, *J. Biol. Chem.* 269: 1-4, 1994). We suggest that it is the regulatory output distinct to the DAF-2 pathway that controls longevity. Alternatively, TGF-β and insulin-like signals may converge only during the L1 stage, when diapause is regulated, and that after this stage, only DAF-2 signaling is necessary for normal metabolic control.

The involvement of insulin and TGF-β signaling in *C. elegans* diapause control suggests that the homologous human pathways may similarly mediate response to famine. Just as environmental extremes can select for variation in the genetic pathways that regulate *C. elegans* dauer formation, famines and droughts in human history may have selected for analogous variants in the human homolog of the *daf* genes. In fact, heterozygous mice carrying either the db or ob recessive diabetes genes, survive fasting about 20% longer than wild type controls (Coleman, *Science* 203: 663-665, 1979). The high frequency of Type II diabetes in many human populations may be the legacy of such selections.

The DAF-3 Smad Protein Anatagonizes DAF-7 TGF-β Receptor Signaling in the C. elegans Dauer Regulatory Pathway

In response to environmental signals *C. elegans* arrests development at the anatomically and metabolically distinctive third-larval dauer stage (Riddle In: C. elegans N, D.L. Riddle, T. Blumenthal, B.J. Meyer, J.R. Priess, eds., *Cold Spring Harbor Press*, 1997, pp. 739-768). Pheromone signal is transduced by chemosensory neurons (Bargmann and Horvitz, *Science* 251:1243, 1991) which couple to a TGF-β signaling pathway (Ren et al., *Science* 274:1389, 1996; Schackwitz et al., *Neuron* 17:719, 1989), as well as an insulin-related signaling pathway (as discussed, *infra*) to trigger changes in the development of the many tissues remodeled in dauer larvae (Riddle, *supra*). Mutations in *daf-7* (a TGF-β homolog (Estevez et al., *Nature* 365:644, 1993)), *daf-4* (a type II TGF-β receptor (Estevez et al., *Nature* 365:644, 1993)), *daf-1* (a type I TGF-β receptor), *daf-8*, and *daf-14* (Smad homolog) cause constitutive arrest at the dauer stage even in the absence of pheromone. These genes constitute a neuroendocrine signaling pathway that is active during non-dauer development: the DAF-7 TGF-β signal is produced by the sensory neuron ASI during nondauer development, whereas *daf-7* expression in this neuron is inhibited during dauer-inducing conditions (Ren, *supra*).

daf-7 and its receptors and Smad proteins are antagonists to *daf-3*. The dauer constitute phenotypes of mutations in the *daf-7* signal transduction pathway genes (including putative null mutations) are fully suppressed by mutations in *daf-3*. These genetic data indicate that in the absence of *daf*-7 signaling, *daf-3* acts to induce dauer arrest.

To discern the molecular basis of the DAF-3 function in this pathway, we determined the sequence and expression pattern of *daf-3*. Cosmids in the *daf-3* genetic region were assayed for gene activity by transformation. Cosmid B0217 partially complemented a *daf-3* mutation, while other cosmids from the

region did not (Fig. 5A). A subclone of B0217 containing only the Smad homolog, but no other coding regions also rescued *daf-3*. Our detection of mutations in the Smad homolog (see below) confirmed its assignment to *daf-3*. Analysis of *daf-3* cDNAs revealed that the gene was transcribed from fifteen exons and was alternatively spliced upstream of the region conserved in Smad proteins. (Fig. 5B) The biological activity of these alternatively spliced isoforms is unknown. The nucleotide (SEQ ID NO: 11) and amino acid sequences (SEQ ID NO: 12) of DAF-3 are shown in Figs. 11 and 12, respectively.

Thus far, the *C. elegans* DAF-3 Smad protein is most closely related in sequence to DPC4, which is a putative cofactor for Smad1, Smad2, and Smad3 (Zhang et al., *Nature*, 383:168, 1996; Lagna et al., *Nature*, 383:832, 1996; Savage et al., *Proc.Natl.Acad.Sci.*, 93:790, 1996; Hahn et al., *Science*, 271:350 (1996). Smads have two conserved domains (Wrana et al., *Trends Genet.*, 12:493, 1996). DAF-3 has these two domains; compared to its closest known relative DPC-4, *daf-3* has 55% amino acid identity in domain I and 30% in domain II (Fig. 5C). However, DPC-4 is not the mammalian DAF-3 homologue: *C. elegans* Sma-4, for example, is more closely related to DPC-4 than DAF-3.

We identified three mutations in *daf-3*, all of which were isolated as suppressors of *daf-7(e1372)*. *mgDf90* is a homozygous viable deletion of 15-90 kb that removes the entire Smad gene (Fig. 5A). *mgDf 90* was identified as a spontaneous mutation that suppressed *daf-7* in the strain of GR1300 (*daf-7 (e1372)* III; *mut-6(st 702) unc-22 (St192) IV*). Thus, suppression of the *daf-7* dauer constitutive phenotype of *daf-3* is *daf-3* null phenotype, demonstrating that wild-type DAF-3 acts antagonistically to signaling from the DAF-7 TGF-β pathway signaling. *daf-3(mg125)* and *daf-3(mg132)* are missense mutations that alter conserved residues in domains 1 and 2 respectively (Fig. 5C). Most

of the mutations detected in other Smads localize to a 45 amino acid segment of domain II (Wrana et al., *Trends in Genet.* 12:493, 1996). Clustering of mutations is observed even in DPC4, for which homozygous null mutations have been identified (Hahn et al., *Science* 271:350, 1996), so the clustering is unlikely to be due to selection for non-null mutations. This hotspot region was sequenced in nine *daf-3* alleles, and no mutations were detected. This difference in mutation location may be a simple statistical anomaly, or may indicate functional differences between DAF-3 and other Smad proteins, consistent with the fact that DAF-3 is antagonized, rather than activated, by an upstream TGF-β molecule.

To determine where DAF-3 may function in control of dauer formation, we examined the expression pattern of a functional *daf-3*/Green Fluorescent Protein (GFP) fusion gene. This was accomplished by replacing a AvrII/SacI fragment from pGP8 with a PCR product in which several restriction sites were inserted after the last codon of *daf-3* before the stop codon. A GFP/*unc*-54 3' end PCR product from pPD95.81 was cloned into the 3' restriction sites to produce pGP19. This DAF-3/GFP fusion partially rescues a *daf-3* mutant (Fig. 7). GFP fluorescence therefore indicates the functional location of DAF-3. DAF-7 signaling from the ASI neuron begins during the L1 stage, and neuron ablations and dauer-formation assays in various environmental conditions indicate that the signal for dauer formation is also received during the first two larval stages (Ren et al., *Science* 274:1389, 1996, Schackwitz et al., *Neuron* 17:719, 1996; Bargmann and Horvitz, *Science* 251:1243, 1991; Golden and Riddle, *Developmental Biology* 102:368, 1984; Swanson and Riddle, *Developmental Biology* 84:27, 1981). Therefore, we most extensively examined L1 larvae.

Almost every transgenic animal showed strong *daf-3*/GFP expression in head neurons (Fig. 6A), the ventral nerve cord (both cell bodies and processes,

see Fig. 6B), the intestinal cells (Fig. 6C), especially the membrane adjacent to the intestinal lumen, the tail hypodermis, and tail neurons. For all GFP scoring, animals were grown at 25-26°C. For scoring of DAF-3/GFP in wild-type and in dauer constitutive mutant backgrounds, three or more lines were scored in each case. A large number of animals were surveyed to determine the expression pattern, and at least 30 animals were scored head-to-tail, and expression was tallied for each tissue. About half of the transgenic animals have weak expression in V blast cells, P blast cells, hyp7 hypodermal cells, and the pharynx. The weak expression impedes cell identification, but the main body of the pharynx is filled, implying expression in pharyngeal muscle (Fig. 6A). Expression is rarely detected in dorsal body wall muscle. The expression pattern in older larvae and adults is similar to that of L1 animals. In addition, DAF-3/GFP is expressed in the distal tip cells and in their precursors, Z1.a and Z4.p, throughout development (Fig. 6D, Fig. 8). DAF-3/GFP is also strongly expressed in unidentified vulval cells. In wild-type embryos of 200-400 cells, DAF-3/GFP is expressed uniformly thoughout the embryo (Fig. 6E). Under the conditions of the experiment, which promote reproductive growth, the subcellular localization of the DAF-3/GFP protein is mainly cytoplasmic (Fig. 6B-E, and see below).

Because DAF-3 activity may be regulated by the DAF-1 and DAF-4 TGF-β receptors, we examined the expression of a DAF-4/GFP fusion in wild-type (Figs. 6A-6G). This construct complements a *daf-4* mutant. A 10 kb SalI fragment from cosmid CO5D2 contains 3 kb of sequence upstream of the *daf-4* transcriptional start, and all of the *daf-4* coding region except codons for the last fourteen residues of *daf-4*. This fragment was subcloned into the SalI site of the GFP plasmid TU#61 (Chalfie et al., *Science* 263: 802-805, 1994). This plasmid was injected into the *daf-4*(*m72*) strain to test the fusion for DAF-4 activity. More than 95% of the transgenic animals were rescued for the dauer-

constitutive and small phenotypes of *daf-4(m72)*, indicating that the fusion has robust DAF-4 activity. The pattern of DAF-4/GFP expression is similar to that of *daf-3*/GFP, except that DAF-4/GFP is localized to membranes, consistent with its role as a receptor. DAF-4/GFP is expressed more strongly in the pharynx (Figs. 6F-G), and more weakly in the ventral nerve cord cell bodies and the body hypodermis. Expression of DAF-4/GFP in wild-type animals is detected later than DAF-3/GFP. DAF-4/GFP is first detectable at late embryogenesis when the embryo resembles an L1 larva. The DAF-4/GFP construct contains an older version of GFP than in DAF-3/GFP; in the older version, the chromophore takes longer to mature. To verify that the difference in embryonic expression of DAF-4/GFP and DAF-3/GFP is not an artefact of the slower maturation time in the *daf-4* strain, we used anti-GFP antibodies to assay GFP. These antibodies should recognize the two forms of GFP equally well. We found that the antibodies recapitulated the results with direct GFP fluorescence: DAF-3/GFP is expressed in early embryos; DAF-4/GFP is not. DAF-4/GFP is also not expressed in membrane surrounding the intestinal lumen, unlike DAF-3/GFP.

The combination of the DAF-3 and DAF-4 expression patterns suggests that these genes act in target tissues to transduce pheromone-regulated DAF-7 neuroendocrine signals. The early expression of DAF-3 in embryos is also consistent with a model that DAF-*3* acts during embryonic development, for example, to mediate the development of neuronal pathways that emit neuroendocrine signals that antagonize DAF-7 TGF-β signaling during the L1 stage. However our data indicates that DAF-3 functions in transducing environmental signals during the L1 and L2 stages. This is supported by the following observations. (1) DAF-7 TGF-β signal from ASI neurons occurs during the L1 and L2 stages and is repressed by dauer-inducing environmental conditions. (2) Expression of the DAF-4 type II receptor begins in very late

embryogenesis. (3) Expression patterns of DAF-3 and DAF-4 are coincident in most of the tissues remodeled during dauer morphogenesis. For example, the cuticle secreted by the hypodermis is modified, the pharynx is slimmed, and the lumen of the intestine is less convoluted. In addition, somatic gonad development is arrested in dauers, and the distal tip cell, in which DAF-3 is expressed, is an important regulator of that development (Kimble, *Developmental Biology* 87:286, 1981). In addition, the intestine and hypodermis of dauer larvae contain large fat stores indicative of a metabolic shift to fat storage. The expression of both the DAF-4 TGF-β family receptor kinase and the DAF-3 Smad protein in these target tissues is consistent with a model that the DAF-7 neuroendocrine signal from the ASI neuron is received directly by these tissues during non dauer development. In addition, the observation that DAF-4 and DAF-3 are expressed in many of the same cells is consistent with a model that DAF-4 signaling to downstream Smads (DAF-8 and DAF-14 are likely candidates) directly regulates DAF-3 gene activity. The TGF-β regulated nuclear localization and transcriptional activation of some Smad proteins suggests that DAF-3 might induce the dauer-specific changes by activating transcription in target tissues of genes required for dauer formation or repressing transcription of genes necessary for nondauer growth.

Smad1 and Smad2 relocalize to become predominantly nuclear when the upstream TGF-β signaling pathways are activated (Baker and Harland, *Genes and Development* 10: 1880, 1996; Hoodless et al., *Cell* 85:489, 1996; Liu et al., *Nature* 381:620, 1996; Macias-Silva et al., *Cell* 87:1215, 1996). In wild-type, DAF-3/GFP is primarily, although not exclusively, cytoplasmic. DAF-3/GFP subcellular distribution was examined in head neurons in the vicinity of ASI (the cell that produces the DAF-7 signal), as well as in intestinal cells. DAF-3/GFP was predominantly cytoplasmic in all animals. However, in all animals, dim GFP fluorescence was observed in the nucleus of some of the

cells with bright fluorescence, and in approximately twenty-five percent of the animals, equivalent DAF-3/GFP levels in the nucleus and cytoplasm has observed in one or more cells.

Because DAF-3 is antagonized by the other members of the DAF-7 TGF-β pathway, we expect that DAF-3 is active (and perhaps localized to the nucleus) when these genes are inactive. We therefore observed the subcellular localization of the full-length DAF-3/GFP fusion protein in the head neurons, tail neurons, and intestine of dauer-constitutive mutant L1 worms, when DAF-3 gene activity is predicted to be highest. In DAF-1(*m402*), *daf-4*(*m72*), *daf-7*(*m62*), *daf-8*(*sa233*), and *daf-14*(*m77*) mutants, DAF-3/GFP was predominantly cytoplasmic, although, as in wild-type, cells were seen with some GFP in the nucleus. In three *daf-4*(*m72*) mutant lines, DAF-3/GFP was localized to the nucleus more than in wild-type lines. When these strains were crossed to wild-type, the increased nuclear localization was seen in both the *daf-4* and wild-type segregants. Thus the increased nuclear GFP was a property of the array, rather than of *daf-4*. Even in the neurons nearest to ASI, where the DAF-7 signal should be strongest, no change in DAF-3/GFP subcellular localization was detected. The DAF-3/GFP fusion protein is predominantly cytoplasmic in L1 and L2 stages of larvae induced to form dauers by environmental conditions or by mutations in the insulin receptor pathway gene *daf-2*, rather than by mutations in the DAF-7 signaling pathway mutants (data not shown). The tissue-specific expression pattern of DAF-3/GFP was unaltered in these mutant backgrounds (data not shown).

The finding that DAF-3/GFP subcellular localization is not strongly responsive to DAF-7 signaling defects or to dauer-inducing environmental conditions does not rule out a role for DAF-3 in the nucleus in dauer formation. Even though we detect no change in DAF-3/GFP subcellular localization, we do detect some DAF-3/GFP in nuclei, and a minor change in nuclear

localization or a change in activity due to phosphorylation state may couple DAF-3 to DAF-7 signaling. In fact, the subcellular localization of *Drosophila* MAD protein is not detectably altered in wild-type when receptor signaling to MAD occurs; relocalization is seen only if the DPP ligand is drastically overexpressed. It is unlikely that a set of undiscovered TGF-β receptors regulates DAF-3. The *C. elegans* genome sequence is 90% complete, and there is only one candidate TGF-β receptor gene other than *daf-1* and *daf-4*. If this receptor were a positive regulator of DAF-3, mutants would be expected to, like *daf-3* mutants, suppress *daf-7* mutants. This receptor acts in a signaling pathway distinct from DAF-3, and it is not a suppressor of *daf-7*.

The implication from Smad homology that DAF-3 is active in the nucleus is supported by two additional observations. First, DAF-3/GFP is associated with chromosomes in intestinal cells during mitosis. These cells divide at the end of the L1 stage, and antibody staining with anti-GFP antibodies and anti-α-tubulin antibodies reveals that DAF-3/GFP is found associated with DNA between the spindles during mitosis (Fig. 8A). We see DAF-3 GFP co-localized with DAPI from prophase to late anaphase. DAF-3/GFP was associated with nuclei in prophase by the following criteria. The spindles were present on either side of the nucleus, but the nucleus has not completely broken down. In particular, an indistinct nucleolus was present. DAF-3/GFP continues to co-localize with DAPI until the chromosomes have separated to the normal distance by which nuclei are separated in the intestine, implying continued association until telophase. At this point in mitosis, DAF-3/GFP fades and becomes undectectable before the nuclei reform the nuclear envelope and nucleolus. Thus, DAF-3 can, indirectly or directly, bind DNA, consistent with the hypothesis that it is a transcriptional activator that acts in the nucleus. DAF-3 is not predicted from its mutant phenotype to have a role

in mitosis. It is possible that the brighter GFP on mitotic chromosomes is due to increased access to DNA due to the breakdown of the nuclear envelope. The second indication of DAF-3 function in the nucleus is our examination of a truncated DAF-3/GFP fusion that is missing most of conserved domain II. The truncated construct pGP7 consists of 8 kb of *daf-3* fused to GFP. An 8 kb EcoR1 fragment from B0217 was cloned into the EcoR1 site of pBluescript SK(-). A PvuI/SalI fragment of this subclone was ligated to a PvuI/SalI fragment from the GFP vector pPD95.81. The resulting plasmid contains ~2.5 kb of sequence upstream of the 5'-most exon of *daf-3* and coding region through the first 58 amino acid residues of domain II. The remaining 175 amino acids of *daf-3* and the 3' noncoding region are replaced with GFP and the *unc-54* 3' end. Three transgenic lines were isolated, and all had a similar phenotype. This fusion protein interferes with dauer induction; like a *daf-3* loss-of-function mutant, it suppresses mutations in *daf-7* (Fig. 7). This truncated protein is predominantly nuclear, suggesting that it represses dauer formation by acting in the nucleus (Fig. 8B). This result implies that wild-type DAF-3 also has a function in the nucleus. The full-length DAF-3/GFP construct also suppresses mutations in *daf-7*, as does a full-length DAF-3 construct without GFP (Fig. 7). This suppression indicates that overexpression of DAF-3 in the cytoplasm has dominant-negative activity, perhaps due to interference with DAF-3 interactions with receptors or cofactors such as other Smads.

The constitutive nuclear localization of truncated DAF-3/GFP fusion gene missing part of domain II suggests that control of Smad localization is complex. A Smad2 construct containing only the conserved domain II of the protein is constitutively nuclear, leading to the suggestion that the C-terminus is an effector domain, and the N-terminus tethers the protein in the cytoplasm (Baker and Harland, *Genes and Development* 10:1880, 1996; Hoodless et al.,

Cell 85:489, 1996; Liu et al., *Nature* 381:620, 1996; and Macias-Silva et al., *Cell* 87:1215, 1996). Our construct, in which the N-terminus is intact, is nuclear. Perhaps both domains provide tethering in the cytoplasm, and any disruption leads to nuclear entry. Alternatively, entry may be differently regulated for DAF-3 and Smad2. Significantly, Smad2, like Smad1 and Smad3 has an SSXS motif at the C terminus (Zhang et al., *Nature* 383:168, 1996; Lagna et al., *Nature* 383:832, 1996; Savage et al., *PNAS* 93:790; Baker and Harland, *Genes and Development* 10:1880, 1996; Hoodless et al., *Cell* 85:489, 1996; Liu et al., *Nature* 381:620, 1996; Macias-Silva et al., *Cell* 87:1215, 1996; and Graf et al., *Cell* 85:479, 1996); this motif is a substrate for phosphorylation and required for nuclear localization of Smad2 (Baker and Harland, *Genes and Development* 10:1880, 1996; Hoodless et al., *Cell* 85:489, 1996; Liu et al., *Nature* 381:620, 1996; and Macias-Silva et al., *Cell* 87:1215, 1996). DAF-3 has a single serine in the C terminal region, and DPC4 has no serines at this location.

We propose a model for the TGF-β pathway in dauer formation (Figs. 9A-B). The DAF-7 TGF-β ligand, which is produced by the ASI sensory neuron in conditions that induce reproductive organ (Ren et al., *Science* 274:1389, 1996; Schakwitz et al., *Neuron* 17:719, 1996), binds to the DAF-1/DAF-4 receptor kinases on target tissues. These receptor kinases then phosphorylate the Smads DAF-8 and/or DAF-14, analogous to the phosphorylation and activation of Smad1, Smad2, and Smad3 (Zhang et al., *Nature* 383:168, 1996; Lagna et al., *Nature* 383:832, 1996; Savage et al., *PNAS* 93:790, 1996). We propose that DAF-3 functions like its closest homolog, DPC4, which dimerizes with phosphorylated Smad1 and Smad2, even under conditions that do not lead to detectable DPC4 phosphorylation (Zhang et al., *Nature* 383:168, 1996; Lagna et al., *Nature* 383:832, 1996; and Savage et al., *PNAS* 93:790). We suggest that DAF-3 forms dauer-inducing homodimers in

the absence of DAF-7 signaling (Figs. 9A-B) that are disrupted when DAF-3 heterodimerizes with a phosphorylated DAF-8 and/or DAF-14 (Fig. 9B). Because *daf-8* and *daf-14* are only partially redundant (Riddle et al., *Nature* 290:668, 1981; Vowels and Thomas, *Genetics* 130:105, 1992; and Thomas et al., *Genetics* 134:1105, 1993), each is likely to perform a unique function in dauer formation. Thus, DAF-3/DAF-8 dimers are proposed to have different activity from DAF-3/DAF-14. Perhaps each activates a subset of genes required for dauer formation. The formation of DAF-8/DAF-3 and/or DAF-14/DAF-3 heterodimers antagonizes dauer induction by the DAF-3/DAF-3 homodimer. A *daf-8(sa233)*; *daf-14(m77)*; *daf-3(mgDf90)* triple mutant can form some dauers in dauer-inducing conditions (data not shown); we suggest that activity of the Daf-2 pathway may induce dauer in this mutant background.

The dauer genetic pathway represents a neuroendocrine pathway for control of a diapause arrest and its associated shifts in metabolism and rates of senescence (Ren et al., *Science* 274:1389, 1996; Schackwitz et al., *Neuron* 17:719, 1996; and Georgi et al., *Cell* 61:635, 1990). Similarly, activins, members of the TGF-β family, were originally identified based on their neuroendocrine regulatory activity, for example, in regulation of gonadotropin signaling (Vale et al., in *Peptide Growth Factors and Their Receptors*, Sporn and Roberts, Eds., Springer-Verlag, Heidelberg, 1990). The DAF-7 signal is not the only signal that is necessary for reproductive development. Because mutations in the DAF-7 TGF-β pathway and in the DAF-2 insulin-like signaling pathway cause the same dauer arrest phenotypes, we propose that both the DAF-7 TGF-β signals and the DAF-2 insulin-like signals are necessary for reproductive development. The involvement of an insulin-like signaling pathway in diapause with its associated metabolic shifts is consistent with metabolic regulation by insulin in vertebrates. Genetic experiments indicate that these pathways act in parallel (Riddle et al., *Nature* 290:668, 1981;

Vowels and Thomas, *Genetics* 130:105, 1992; and Thomas et al., *Genetics* 134:1105, 1993). In particular, *daf-3* mutants efficiently suppress *daf-7* mutants, but not *daf-2* mutants, and *daf-16* mutants efficiently suppress *daf-2* mutants, but poorly suppress *daf-7* mutants. It is not yet clear whether these two signaling pathways coverage on target tissues or in other regulatory (e.g., hormone secreting) cells. However, the expression of the DAF-7 receptor pathway genes and the DAF-16 gene in essentially all target tissues suggests that the TGF-β and insulin pathways act there, and therefore that integration must occur there. Thus, we suggest in Figs. 9A and 9B that the DAF-2 pathway converges on DAF-3/DAF-8DAF-1 Smad signaling to regulate metabolic gene expression in target tissues.

The integration of insulin-like and TGF-β signals in metabolic control has important implications for the molecular basis of diabetes. For example, these converging pathways for dauer control suggest that in human metabolic control both a DAF-7-like signal and insulin may be necessary for full metabolic control. Thus, declines in signaling from the human homolog of DAF-7 could underlie the insulin resistance associated with Type II diabetes. In fact the dauer pheromone has been reported to be a fatty acid and to cause down-regulation of DAF-7 expression (Ren et al., supra). Thus pheromone regulation of metabolism may be related to mammalian obesity induced diabetes, and a human mutation in DAF-7 or its receptors is expected to contribute to a diabetic condition, just like mutations in the insulin receptor. In addition if obesity or age or both cause human DAF-7 to decline, e.g., under high leptin conditions, such a result would explain late onset/obesity related diabetes.

Cloning Mammalian DAF Sequences

Based on our isolation of novel nematode DAF cDNAs, the isolation of

mammalian DAF nucleic acid sequences, including human DAF sequences, is made possible using the sequences described herein and standard techniques. In particular, using all or a portion of a nematode DAF sequence, one may readily design oligonucleotide probes, including degenerate oligonucleotide probes (i.e., a mixture of all possible coding sequences for a given amino acid sequence). These oligonucleotides may be based upon the sequence of either strand of the DNA.

Exemplary probes or primers for isolating mammalian DAF sequences preferably correspond to conserved blocks of amino acids, for example, conserved DAF motifs. Exemplary motifs are as follows:

DAF-2 (tyrosine kinase domain) (SEQ ID NO: 33)
1242 KFHEWAAQICDGMAYLESLKFCHRDLAARNCMINRDETVKIGDFGM
ARDLFYHDYYKPSGKRMMPVRWMSPESLKDGKFDSKSDVWSFGVVLYE
MVTLGAQPYIGLSNDEVLNYIGMARKVIKKPEC 1368

DAF-2 (ligand binding domain) (SEQ ID NO: 34)
242 NTTCQKSCAYDRLLPTKEIGPGCDANGDRCHDQCVGGCERVNDATA
CHACKNVYHKGKCIEKCDAHLYLLLQRRCVTREQCLQLNPVLSNKTVPIK
ATAGLCSDKCPDGYQINPDDHRECRKCVGKCEIVC 372

DAF-2 (67 amino acid motif) (SEQ ID NO: 79)
1158 AIKINVDDPASTENLNYLMEANIMKNFKTNFIVQLYGVISTVQPAMV
VMEMMDLGNLRDYLRSKRED 1224

DAF-2 (54 amino acid motif) (SEQ ID NO: 80)
1362 VIKKPECCENYWYKVMKMCWRYSPRDRPTFLQLVHLLAAEASPEFR
DLSFVLTD 1415

DAF-2 (69 amino acid motif) (SEQ ID NO: 81)

404 KQDSGMASELKDIFANIHTITGYLLVRQSSPFISLNMFRNLRRIEAKSL
FRNLYAITVFENPNLKKLFD 472

DAF-2 (52 amino acid motif) (SEQ ID NO: 82)
98 FPHLREITGTLLVFETEGLVDLRKIFPNLRVIGGRSLIQHYALIIYRN
PDLE 149

DAF-2 (46 amino acid motif) (SEQ ID NO: 83)
149 EIGLDKLSVIRNGGVRIIDNRKLCYTKTIDWKHLITSSINDVVVDN 194
DAF-2 (36 amino acid motif) (SEQ ID NO: 84)
1112 YNADIDWELRQDDVVLGQQCGEGSFGKVYLGTGNNVV 1147

DAF-3 (Smad Domain I) (SEQ ID NO: 35)
240 FDQKACESLVKKLKDKKNDLQNLIDVVLSKGTKYTGCITIPRTLDGR
LQVHGRKGFPHVVYGKLWRFNEMTKNETRHVDHCKHAFEMKSDMVC
VNPYHYEIVI 342

DAF-3 (Smad Domain II) (SEQ ID NO: 36)
690 NRYSLGLEPNPIREPVAFKVRKAIVDGIRFSYKKDGSVWLQNRMKYPV
FVTSGYLDEQSGGLKKDKVIIKVYGCASIKTF 768

DAF-3 (79 amino acid motif) (SEQ ID NO: 85)
819 DSLAKYCCVRVSFCKGFGEAYPER 842

DAF-16 (forkhead DNA binding domain) (SEQ ID NO: 37)
727 KKTTTRRNAWGNMSYAELITTAIMASPEKRLTLAQVYEWMVQNVPY
FRDKGDSNSSAGWKNSIRHNLSLHSRFMRIQNEGAGKSSWWVINPDAKPG
MNPRRTRERS 1044

DAF-16 (103 amino acid motif) (SEQ ID NO: 54)
242 KKTTTRRNAWGNMSYAELITTAIMASPEKRLTLAQVYEWMVQNVPY

FRDKGDSNSSAGWKNSIRHNLSLHSRFMRIQNEGAGKSSWWVINPDAKPG
MNPRRTR 344

DAF-16 (41 amino acid motif) (SEQ ID NO: 55)
137 TFMNTPDDVMMNDDMEPIPRDRCNTWPMRRPQLEPPLNSSP 177

DAF-16 (109 amino acid motif) (SEQ ID NO: 56)
236 DDTVSGKKTTTRRNAWGNMSYAELITTAIMASPEKRLTLAQVYEWM
VQNVPYFRDKGDSNSSAGWKNSIRHNLSLHSRFMRIQNEGAGKSSWWVI
NPDAKPGMNPRRTR 344

DAF-16 (98 amino acid motif) (SEQ ID NO: 58)
372 KPNPWGEESYSDIIAKALESAPDGRLKLNEIYQWFSDNIPYFGERSSPE
EAAGWKNSIRHNLSLHSRFMRIQNEGAGKSSWWVINPDAKPGMNP
RRTR 469

Using such motifs, mammalian DAF-2, DAF-3, and DAF-16 genes may be isolated from sequence databases (for example, by the use of standard programs such as Pileup). Alternatively, such sequences may be used to design degenerate oligonucleotide probes to probe large genomic or cDNA libraries directly. General methods for designing and preparing such probes are provided, for example, in Ausubel et al., *Current Protocols in Molecular Biology*, 1996, Wiley & Sons, New York, NY; and *Guide to Molecular Cloning Techniques*, 1987, S. L. Berger and A. R. Kimmel, eds., Academic Press, New York. These oligonucleotides are useful for DAF gene isolation, either through their use as probes for hybridizing to DAF complementary sequences or as primers for various polymerase chain reaction (PCR) cloning strategies. If a PCR approach is utilized, the primers are optionally designed to allow cloning of the amplified product into a suitable vector. PCR is particularly useful for screening cDNA libraries from rare tissue types.

Hybridization techniques and procedures are well known to those skilled in the art and are described, for example, in Ausubel et al., *supra*, and *Guide to Molecular Cloning Techniques*, *supra*. If desired, a combination of different oligonucleotide probes may be used for the screening of the recombinant DNA library. The oligonucleotides are, for example, labelled with ^{32}P using methods known in the art, and the detectably-labelled oligonucleotides are used to probe filter replicas from a recombinant DNA library. Recombinant DNA libraries (for example, human cDNA libraries, such as hypothalamus- or pancreas-derived cDNA libraries, particularly for DAF-2 and DAF-7 cDNAs) may be prepared according to methods well known in the art, for example, as described in Ausubel et al., *supra*, or may be obtained from commercial sources.

For detection or isolation of closely related DAF sequences, high stringency hybridization conditions may be employed; such conditions include hybridization at about 42°C and about 50% formamide; a first wash at about 65°C, about 2X SSC, and 1% SDS; followed by a second wash at about 65°C and about 0.1% SDS, 1X SSC. Lower stringency conditions for detecting DAF genes having less sequence identity to the nematode DAF genes described herein include, for example, hybridization at about 42°C in the absence of formamide; a first wash at about 42°C, about 6X SSC, and about 1% SDS; and a second wash at about 50°C, about 6X SSC, and about 1% SDS.

As discussed above, DAF-specific oligonucleotides may also be used as primers in PCR cloning strategies. Such PCR methods are well known in the art and are described, for example, in *PCR Technology*, H.A. Erlich, ed., Stockton Press, London, 1989; *PCR Protocols: A Guide to Methods and Applications*, M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White, eds., Academic Press, Inc., New York, 1990; and Ausubel et al., *supra*. Again, sequences corresponding to conserved regions in a DAF sequence (for example, those regions described above) are preferred for use in isolating

mammalian DAF sequences. Such probes may be used to screen cDNA as well as genomic DNA libraries.

Sequences obtained are then examined (for example, using the Pileup program) to identify those sequences having the highest amino acid sequence identity to the *C. elegans* sequence, particularly in or between conserved DAF domains (for example, those domains described above). In one particular example, the human FKHR and AFX genes are 10^{33} more closely related to the DAF-16 forkhead domain than the next most closely related forkhead domain protein, making FKHR and AFX candidates for mammalian DAF-16 genes.

Following isolation of such candidate genes by sequence homology, the genes are then tested for their ability to functionally complement a *daf* mutation. This is most readily assayed by transformation of the sequence into a *C. elegans* strain having an appropriate mutant background. Exemplary *C. elegans* transformation techniques are described, for example, in Mello et al., *EMBO J.* 10: 3959-3970, 1991, and assays for DAF-2, DAF-3, and DAF-16 polypeptide function are described herein. To be considered useful in the invention, a mammalian sequence need not fully complement a *C. elegans* defect, but must provide a detectable level of functional complementation.

The DAF, AGE, or AKT gene homologue identified as above, may also complement or alter the metabolic phenotypes of a mammalian cell line.

For example, addition of DAF-7, TGF-β-like growth factor to an insulin responsive cell line (e.g., the 3T3-L1 cell line) may accentuate insulin responsiveness. Similarly genetic transformation of such a cell line with wild type or dominantly activated versions of a DAF, AGE, or AKT gene may alter metabolism. Such perturbations of metabolic control are stringent tests of candidate genes as DAF, AGE, or AKT homologues.

In addition, if that mammalian candidate homologue acts in a metabolic control pathway, and is expressed in similar metabolic control tissues (liver,

adipose), it is likely to function homologously to DAF proteins from *C. elegans*.

Addition of a wild type or activated DAF, AKT, or AGE protein (for example by VP16 activation of the DAF-3 or DAF-16 transcription factors) can confer on cell lines altered metabolic phenotypes. Thus supplying *daf, age,* or *akt* gene activity to such a cell line can alter its metabolism. This is one explemplary test of homologous DAF function in metabolic control.

DAF Polypeptide Expression

In general, DAF polypeptides according to the invention may be produced by transformation of a suitable host cell with all or part of DAF-encoding cDNA fragment (e.g., one of the cDNAs described herein or isolated as described above) in a suitable expression vehicle.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the recombinant protein. The precise host cell used is not critical to the invention. The DAF polypeptide may be produced in a prokaryotic host (e.g., E. coli) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae*, insect cells, e.g., Sf9 or Sf21 cells, or mammalian cells, e.g., COS 1, NIH 3T3, or HeLa cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD; also, see, e.g., Ausubel et al., *supra*). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel et al. (*supra*); expression vehicles may be chosen from those provided, e.g., in Cloning Vectors: A Laboratory Manual (P.H. Pouwels et al., 1985, Supp. 1987).

One preferred expression system is the baculovirus system (using, for example, Sf9 cells and the method of Ausubel et al., *supra*). Another

baculovirus system makes use of the vector pBacPAK9 and is available from Clontech (Palo Alto, CA).

Alternatively, an DAF polypeptide is produced in a mammalian system, for example, by a stably-transfected mammalian cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public, e.g., see Pouwels et al. (*supra*); methods for constructing such cell lines are also publicly available, e.g., in Ausubel et al. (*supra*). In one example, cDNA encoding the DAF protein is cloned into an expression vector which includes the dihydrofolate reductase (DHFR) gene. Integration of the plasmid and, therefore, the DAF protein-encoding gene into the host cell chromosome is selected for by inclusion of 0.01-300 μM methotrexate in the cell culture medium (as described in Ausubel et al., *supra*). This dominant selection may be accomplished in most cell types. Recombinant protein expression may be increased by DHFR-mediated amplification of the transfected gene. Methods for selecting cell lines bearing gene amplifications are described in Ausubel et al. (*supra*); such methods generally involve extended culture in medium containing gradually increasing levels of methotrexate. DHFR-containing expression vectors commonly used for this purpose include pCVSEII-DHFR and pAdD26SV(A) (described in Ausubel et al., supra). Any of the host cells described above or, preferably, a DHFR-deficient CHO cell line (e.g., CHO DHFR$^-$ cells, ATCC Accession No. CRL 9096) are among the host cells preferred for DHFR selection of a stably-transfected cell line or DHFR-mediated gene amplification.

In yet other alternative approaches, the DAF polypeptide is produced *in vivo* or, preferably, *in vitro* using a T7 system (see, for example, Ausubel et al., *supra*, or other standard techniques).

Once the recombinant DAF protein is expressed, it is isolated, e.g., using affinity chromatography. In one example, an anti-DAF protein antibody (e.g.,

produced as described herein) may be attached to a column and used to isolate the DAF protein. Lysis and fractionation of DAF protein-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., *supra*).

Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, Laboratory Techniques In Biochemistry And Molecular Biology, eds., Work and Burdon, Elsevier, 1980).

Polypeptides of the invention, particularly short DAF polypeptide fragments, may also be produced by chemical synthesis (e.g., by the methods described in Solid Phase Peptide Synthesis, 2nd ed., 1984 The Pierce Chemical Co., Rockford, IL).

These general techniques of polypeptide expression and purification may also be used to produce and isolate useful DAF fragments or analogs (described herein).

Anti-DAF Antibodies

Using any of the DAF polypeptides described herein or isolated as described above, anti-DAF antibodies may be produced by any standard technique. In one particular example, a DAF cDNA or cDNA fragment encoding a conserved DAF domain is fused to GST, and the fusion protein produced in E. coli by standard techniques. The fusion protein is then purified on a glutathione column, also by standard techniques, and is used to immunize rabbits. The antisera obtained is then itself purified on a GST-DAF affinity column, for example, by the method of Finney and Ruvkun (*Cell* **63**:895-905, 1990), and is shown to specifically identify GST-DAF, for example, by Western blotting.

Polypeptides for antibody production may be produced by recombinant

or peptide synthetic techniques (see, e.g., *Solid Phase Peptide Synthesis*, *supra*; Ausubel et al., *supra*).

For polyclonal antisera, the peptides may, if desired, be coupled to a carrier protein, such as KLH as described in Ausubel et al, *supra*. The KLH-peptide is mixed with Freund's adjuvant and injected into guinea pigs, rats, or preferably rabbits. Antibodies may be purified by any method of peptide antigen affinity chromatography.

Alternatively, monoclonal antibodies may be prepared using a DAF polypeptide (or immunogenic fragment or analog) and standard hybridoma technology (see, e.g., Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976; Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., *In Monoclonal Antibodies and T Cell Hybridomas*, Elsevier, NY, 1981; Ausubel et al., *supra*).

Once produced, polyclonal or monoclonal antibodies are tested for specific DAF recognition by Western blot or immunoprecipitation analysis (by the methods described in Ausubel et al., *supra*). Antibodies which specifically recognize a DAF polypeptide described herein are considered to be useful in the invention. Anti-DAF antibodies, as isolated above, may be used, e.g., in an immunoassay to measure or monitor the level of DAF polypeptide produced by a mammal or to screen for compounds which modulate DAF polypeptide production (for example, in the screens described herein). In one particular example, antibodies to human DAF-7 polypeptide are useful for screening blood samples from patients to determine whether they possess decreased DAF-7 polypeptide levels. Such antibodies may be used in any immunological assay, for example, an ELISA assay, and a decrease in DAF-7 is taken as an indication of a diabetic condition, for example, obesity onset Type II diabetes. In another particular example, anti-DAF antibodies are useful for carrying out pedigree analysis. For example, blood samples from individuals may be

screened with anti-DAF-7 antibodies to detect those members of a family with a predisposition to a diabetic condition. Anti-DAF antibodies may also be used to identify cells that express a DAF gene.

DAF-7 therapy for obesity-onset Type II diabetes

Our data indicates that DAF-7 represents an endocrine hormone for metabolic control that acts synergistically with insulin. Declines in DAF-7 may be induced by obesity, just as the dauer pheromone, a fatty acid, causes declines in *C. elegans* DAF-7 production.

Accordingly, obesity onset Type II diabetes, glucose intolerance, and the associated atherosclerosis may be treated if DAF-7 hormone is injected intramuscularly or intravenously (Fig. 23).

In addition, antibodies to human DAF-7 should detect declines in DAF-7 in pre-diabetic, glucose-intolerant, or obesity induced diabetes. Such antibodies will detect DAF-7 levels in blood, just as insulin levels are detected in metabolic disease.

DAF-7 therapeutic potential and dosage can be developed in mouse models of obesity onset diabetes, for example, the db and ob mouse.

DAF-7 may be injected either intravenously or intramuscularly, in analogy to insulin therapy.

The decision of which classes of diabetics to treat with DAF-7 will come from a combination of blood tests for DAF-7 levels and genetic testing to determine which daf, age, or akt mutations a particular diabetic or pre-diabetic patient carries.

Screening Systems for Identifying Therapeutics

Based on our experimental results, we have developed a number of screening procedures for identifying therapeutic compounds (e.g., anti-diabetic

and anti-obesity pharmaceuticals or both) which can be used in human patients. In particular examples, compounds that down regulate *daf-3* or *daf-16* or their human homologs are considered useful in the invention. Similarly, compounds that up regulate or activate *daf-1*, *daf-2*, *daf-4*, *daf-7*, *daf-8*, *daf-11* *daf-14*, *age-1*, and *akt* (or each of their corresponding human homologs) are also considered useful as drugs for the treatment of impaired glucose tolerance conditions, such as diabetes and obesity. In general, the screening methods of the invention involve screening any number of compounds for therapeutically active agents by employing any number of *in vitro* or *in vivo* experimental systems. Exemplary methods useful for the identification of such compounds are detailed below.

The methods of the invention simplify the evaluation, identification, and development of active agents for the treatment and prevention of impaired glucose tolerance conditions, such as diabetes and obesity. In general, the screening methods provide a facile means for selecting natural product extracts or compounds of interest from a large population which are further evaluated and condensed to a few active and selective materials. Constituents of this pool are then purified and evaluated in the methods of the invention to determine their anti-diabetic or anti-obesity activities or both.

Below we describe screening methods for evaluating the efficacy of a compound as anti-diabetic or anti-obesity agents or both. These examples are intended to illustrate, not limit, the scope of the claimed invention.

Test Extracts and Compounds

In general, novel drugs for the treatment of impaired glucose tolerance conditions are identified from large libraries of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the field of drug discovery and

development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the exemplary methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceangraphics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, e.g., by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (e.g., taxonomic dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known for their anti-diabetic and anti-obesity activities should be employed whenever possible.

When a crude extract is found to have anti-diabetic or anti-obesity activities or both, further fractionation of the positive lead extract is necessary

to isolate chemical constituents responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having anti-diabetic or anti-obesity activities. The same *in vivo* and *in vitro* assays described herein for the detection of activities in mixtures of compounds can be used to purify the active component and to test derivatives thereof. Methods of fractionation and purification of such heterogenous extracts are known in the art. If desired, compounds shown to be useful agents for the treatment of pathogenicity are chemically modified according to methods known in the art. Compounds identified as being of therapeutic value are subsequently analyzed using any standard animal model of diabetes or obesity known in the art.

There now follow examples of high-throughput systems useful for evaluating the efficacy of a molecule or compound in treating (or preventing) an impaired glucose tolerance condition.

Nematode Release of Dauer Arrest Bioassays

To enable mass screening of large quantities of natural products, extracts, or test compounds in an efficient and systematic fashion, *C. elegans* mutant dauer larvae (e.g., *C. elegans* containing mutations described herein, such as *C. elegans daf-2* mutant dauer larvae) are cultured in wells of a microtiter plate, facilitating the semiautomation of manipulations and full automation of data collection. As discussed above, compounds that down regulate DAF-3 or DAF-16 activities or up regulate DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT activities are considered useful in the invention. Such compounds are identified by their effect on dauer formation in *C. elegans* strains carrying mutations in these genes (as described above).

In particular examples, nematodes bearing mutations in the DAF-2 polypeptide arrest as dauer larvae, never producing progeny. All of the metabolic and growth arrest phenotypes caused by lack of *daf-2* are suppressed by mutations in *daf-16*. Mutations in the PI 3-kinase, AGE-1, have the same phenotype as lack of *daf-2*, and such mutations are also suppressed by *daf-16* mutations. Biochemical analysis of insulin signaling in mammals supports the view that AGE-1 transduces signals from the DAF-2 receptor by generating a PIP3 signal. Because *daf-16* mutations suppress lack of *daf-2*, or *age-1* gene activity, it is believed that PIP3 down regulates or modifies *daf-16* gene activity. The biochemical overlap between DAF-2/AGE-1 and insulin receptors/PI 3-kinase indicates that the human homolog of the *C. elegans daf-16* gene acts in the insulin pathway as well. Thus, the *C. elegans* insulin signaling pathway yields the surprising result that the animals can live without insulin signaling, provided they are mutant in *daf-16*. This analysis therefore indicates that a compound that inhibits DAF-16 activity would reverse the effects of diabetic lesions, e.g., in the production or secretion of insulin or in the reception of insulin signals by target tissues. Such drugs would be expected to be efficacious in the treatment of insulin deficiencies due to pancreatic β cell destruction in Type I diabetes, as well as some Type II diabetes due to defects in insulin signaling.

To evaluate the ability of a test compound or an extract to decrease *daf-16* gene activity, mutant *daf-2* (e1370); *daf-16* (mgDf50) animals carrying an integrated human DAF-16 gene are incubated in microtiter dishes in the presence of a test compound. This human DAF-16 gene supplies all of the DAF-16 activity in the *C. elegans* strain and thus allows *daf-2*-induced dauer arrest unless its activity is decreased by the candidate test compound. If desired, various concentrations of the test compound or extract can be inoculated to assess the dosage effect. Control wells are incubated in the

absence of a test compound or extract. Plates are then incubated at 25°C. After an appropriate period of time, e.g., 2 to 5 days, wells are examined for progeny. The presence of progeny is taken as an indication that the test compound or extract is effective at inhibiting *daf-3* or *daf-16* activity, and therefore is considered useful in the invention. Any compound that inhibits DAF-16 gene activity (or activates upstream signaling in the absence of receptor function) will allow reproduction. This is shown schematically in Fig. 19.

Alternatively, a diabetic condition may arise from defects in the DAF-7 TGF-β signaling pathway. Since a decrease in DAF-3 activity bypasses the need for DAF-7 activity in *C. elegans* metabolic control, drugs that down regulate DAF-3 activity are useful for ameliorating the metabolic defects associated with diabetes. To screen for such drugs, *daf-7* (e1372); *daf-3* (mg90) nematodes expressing human DAF-3 are exposed to chemicals as described above. In this strain, human DAF-3 supplies all DAF-3 activity, causing *daf-7* induced dauer arrest unless its activity is inhibited (Fig. 20). Compounds capable of inhibiting this activity are considered useful therapeutics in the invention.

Finally, in a less complex screen for drugs that inhibit *C. elegans daf-3* or *daf-16*, *daf-7* or *daf-2* mutants are directly screened for compounds that decrease *C. elegans daf-3* or *daf-16* gene activity.

In addition, *C. elegans* worms carrying other *daf* mutations may be utilized in an assay to obtain additional information on the mode of action of the test compound in the insulin or TGF-β signaling pathways. For example, a drug having PIP3 agonist activity would be expected to allow *age-1* and *daf-2* mutants (but not *akt* or *daf-7* mutants) to not arrest at the dauer stage. Similarly, drugs that inhibit *daf-3* are expected to suppress *daf-7* mutants but not *daf-2* or *age-1* mutants.

Other Screening Assays

Other drug screening assays may also be performed using either *C. elegans* worms or mammalian cell cultures. If desired, such assays may include the use of reporter gene constructs.

For example, evaluation of the effects of test compounds on dauer formation or reporter gene expression in mutant *C. elegans* strains expressing particular human homologs of the *daf, age, or akt* genes (i.e., humanized *C. elegans*) represent useful screening methods. Expression of the human homologs in *C. elegans* is accomplished according to standard methods and, if desired, such genes may be operatively linked to a gene promoter obtained from *C. elegans*. Such promoters include, without limitation, the *C. elegans daf-16*, *age-1*, *daf-3*, *daf-4*, and *akt* gene promoters. For example, the 2.5 kb *age-1* promoter can be generated and isolated by employing standard PCR methods using the following primers:
5'GGAAATATTTTAGGCCAGATGCG3' (SEQ IS NO: 49) and
5'CGGACAGTCCTGAATACACC3' (SEQ ID NO: 50).

Additionally, mammalian tissue culture cells expressing *C. elegans daf*, *age-1*, or *akt* homologs may be used to evaluate the ability of a test compound or extract to modulate the insulin or TGF-β signaling pathways. Because the signaling pathways from the ligands, receptors, kinase cascades, and downstream transcription factors are conserved from man to worm, test compounds or extracts that inhibit or activate the worm signaling proteins should also inhibit or activate their respective human homolog. For example, our identification that DAF-16 is a transcription factor that acts downstream of insulin-like signaling in *C. elegans* indicates that human DAF-16 transcription reporter genes also can be used to identify drugs that inhibit any of the kinases in the signaling pathway downstream of insulin signaling. For example, the use of DAF-16 and DAF-3 protein binding sites in reporter gene constructs may be

used to monitor insulin signaling. Candidate compounds mimicking insulin signaling (e.g., PIP3 agonists) are expected to increase reporter gene expression and are considered useful in the invention.

Reporter Gene Construct

In one particular example, the invention involves the use of a reporter gene that is expressed under the control of a *C. elegans* gene promoter, e.g., a promoter that includes the TCTCGTTGTTTGCCGTCGGATGTCTGCC (SEQ ID NO: 51) enhancer element, such as the *C. elegans* pharyngeal myosin promoter (Okkema and Fire, *Development* 120: 2175-2186, 1994). This enhancer element is known to respond to DAF-3 regulation (i.e., in *daf-7* mutants, where *daf-3* is active, the element confers low level expression to reporter genes; whereas in a *daf-7*; *daf-3* mutant (for example, *daf-7* (e1372); *daf-3*), the element confers low level expression to reporter genes). Other equivalent enhancer elements may also be used in the invention, e.g., the enhancer element which is bound by the *Xenopus* Smad1 and Fast1 forkhead proteins (*Nature* 383 600-608, 1996). The enhancer element is cloned upstream of any standard reporter gene, e.g., the luciferase or green fluorescent protein (GFP) reporter genes. In preferred embodiments, the GFP reporter gene is used in *C. elegans*. In other preferred embodiments, either the GFP or the luciferase reporter genes may used in a mammalian cell based assay. The reporter gene construct is subsequently introduced into an appropriate host (e.g., *C. elegans* or a mammalian cell) according to any standard method known in the art. Analysis of reporter gene activity in the host organism or cell is determined according to any standard method, e.g., those methods described herein. Such reporter gene (and host cell systems) are useful for screening for drugs that modulate insulin or DAF-7 metabolic control signaling.

C. elegans

In one working example, the above-described reporter gene construct is introduced into wild-type *C. elegans* according to standard methods known in the art. If the enhancer element is operational, then it is expected that reporter gene expression is turned on. Alternatively, in *daf* mutants (e.g., *daf-7* or *daf-2* mutants, where insulin signaling is defective) carrying the above-described reporter gene construct, reporter gene activity is turned off.

Using this on/off distinction, test compounds or extracts are evaluated for the ability to disrupt the signaling pathways described herein. In one working example, *daf-2* mutant worms carrying the reporter gene construct are used to assay the expression of the reporter gene. The majority of worms expressing the reporter gene will arrest at the dauer stage because of the *daf-2* phenotype. If however the test compound or extract inhibits DAF-16 activity, then the worms will exhibit a *daf-2*; *daf-16* phenotype (i.e., do not arrest), developing to produce eggs. Such eggs are selected using a bleach treatment and reporter gene expression in the test compound or extract is assayed according to standard methods, e.g., worms are examined with an automated fluorometer to reveal the presence of reporter gene expression, e.g., GFP. Candidate compounds that suppress the *daf-2* phenotype or turn on reporter gene expression, i.e., activate signals in the absence of DAF-2 receptor (e.g., PIP3 mimetics) or inactivate DAF-16, are considered useful in the invention.

Analogous screens may also be performed using *daf-7* mutants as a means to identify drugs that inactivate other *daf*-genes, such as DAF-3, or compounds that activate the DAF-1/DAF-4 receptors. Such screens may be coupled to reporter screens, for example, using GFP reporter genes whose expression is under DAF-3 transcriptional control (e.g., the myoII element). Drugs identified in such screens are useful as DAF-7 mimetics. Because DAF-7 expression may be down regulated in obesity, such drugs are useful in

the treatment of obesity induced

Type II diabetes

In yet another working example, *C. elegans* DAF-3 and DAF-16 genes can be replaced with a human homolog, (e.g., FKHR for DAF-16), and screens similar to those described above performed in the nematode system. Because drugs may act upstream of the transcription factors, it is useful to replace DAF-1, DAF-4, DAF-8,DAF-14, DAF-2, DAF-3, DAF-16, or AGE-1 with the appropriate human homolog, and to screen the humanized *C. elegans* animals. Such screens are useful for identifying compounds having activities in humans.

Mammalian Cells

Mammalian insulin-responsive cell lines are also useful in the screening methods of the invention. Here reporter gene constructs (for example, those described above) are introduced into the cell line to evaluate the ability of a test compound or extract to modulate insulin and TGF-β signaling pathways using a second construct expressing a *C. elegans daf, age, or akt* gene or their corresponding human homologs. Exemplary cell lines include, but are not limited to, mouse 3T3, L6 ,and L1 cells (MacDougald et al., *Ann. Rev. Biochem.* 64: 345-373, 1995) Introduction of the constructs into such cell lines is carried out according to standard methods well known in the art.

To test a compound or extract, it is added to the cell line, and reporter gene expression is monitored. Compounds that induce reporter gene expression in the absense of insulin or DAF-7 signaling are considered useful in the invention. Such compounds may also turn the cells into adipocytes, as insulin does.

Compounds identified in mammalian cells may be tested in other screening assays described herein, and, in general, test compounds may be

assayed in multiple screens to confirm involvement in insulin or DAF-7 signaling.

Metabolic control by DAF-7 protein may be tested using any known cell line, e.g., those described herein.

In Vitro Screening Methods

A variety of methods are also available for identifying useful compounds in *in vitro* assays. In one particular example, test compounds are screened for the ability to activate the phosphorylation of Smad proteins, DAF-8, DAF-14, or
DAF-3, by DAF-1 or DAF-4 *in vitro*. In these assays, DAF-8, DAF-14, or DAF-3 is preferably tagged with a heterologous protein domain, for example, the myc epitope tag domain(s) by the method of Ausubel et al., and are incubated with the C-terminal kinase domain of DAF-1 or DAF-4. Phosphorylation of the Smad proteins is preferably detected by immunoprecipitation using antibodies specific to the tag, followed by scintillation counting. Test compounds may be screened in high throughout microtiter plate assays. A test compound that effectively stimulates the phosphorylation of DAF-8, DAF-14, or DAF-3 is considered useful in the invention. Using these same general assays, compounds that activate the phosphorylation of DAF-16 by AKT or GSK-3 may also be identified.

In another working example, test compounds are screened for the ability to inhibit the *in vitro* association of DAF-16 and the Smad proteins DAF-3, or preferentially activates the association of DAF-16 with DAF-8 and DAF-14, DAF-8, or DAF-14, or to inhibit the association of DAF-3 and DAF-16 with DNA *in vitro*. These assays are carried out by any standard biochemical methods that test protein-protein binding or protein-DNA binding. In one particular example, to test for interactions between proteins, each protein is

tagged with a different heterologous protein domain (as described above). Immunoprecipitations are carried out using an antibody to one tag, and an ELISA assay is carried out on the immunoprecipitation complex to test for the presence of the second tag. In addition, if interaction capability is enhanced by a DAF or AKT kinase, this protein is also preferably included in the reaction mixture. Similarly, to test for interactions of these proteins with DNA, antibodies to the tag are utilized in immunoprecipitations, and the presence of the DNA detected by the presence of the DNA label in the immunoprecipitation complex. A test compound that effectively inhibits the association between these proteins or between DAF-3 and DAF-16 with DNA or both is considered useful in the invention.

In still another working example, test derivatives of PIP3 are screened for the ability to increase *in vitro* AKT activity. This is accomplished, in general, by combining a labeled PIP3 and an AKT polypeptide in the presence and absence of the test compound under conditions that allow PIP3:AKT to bind *in vitro*. Compounds are then identified that interfere with the formation of the PIP3:AKT complex. Test compounds that pass this first screen may then be tested for increased AKT activation *in vitro* using GSK3 targets, or may be tested in nematodes or mammalian cells (as described above). An increase in AKT kinase activity is taken as an indication of a compound useful for ameliorating or delaying an impaired glucose tolerance condition.

In yet another working example, DAF-3 or DAF-16 may be expressed in a yeast one-hybrid assay for transcriptional activation. Methods for such assays are described in Brent and Ptashne (*Cell* 43:729-736, 1985). A test compound that blocks the ability of DAF-3 or DAF-16 or both to activate (or repress) transcription in this system is considered useful in the invention.

In a final working example, compounds may be screened for their ability to inhibit an interaction between any of DAF-3, DAF-8, and DAF-14, or

between DAF-3 and DAF-16. These *in vivo* assays may be carried out by any "two-hybrid" or "interaction trap" method (for example, by using the methods described by Vijaychander et al (*Biotechniques* 20: 564-568)).

Modulatory Compounds

Our experimental results facilitate the isolation of compounds useful in the treatment of impaired glucose tolerance diseases that are antagonists or agonists of the insulin or TGF-β signaling pathways identified in *C. elegans* and described above. Exemplary methods for the isolation of such compounds now follow.

Antagonists

As discussed above, useful therapeutic compounds include those which down regulate the expression or activity of DAF-3 or DAF-16. To isolate such compounds, DAF-3 or DAF-16 expression is measured following the addition of candidate antagonist molecules to a culture medium of DAF-3 or DAF-16-expressing cells. Alternatively, the candidate antagonists may be directly administered to animals (for example, nematodes or mice) and used to screen for their effects on DAF-3 or DAF-16 expression.

DAF-3 or DAF-16 expression is measured, for example, by standard Northern blot analysis (Ausubel et al., *supra*) using a DAF-3 or DAF-16 nucleic acid sequence (or fragment thereof) as a hybridization probe. The level of DAF-3 or DAF-16 expression in the presence of the candidate molecule is compared to the level measured for the same cells, in the same culture medium, or in a parallel set of test animals, but in the absence of the candidate molecule. Preferred modulators for anti-diabetic or anti-obesity purposes are those which cause a decrease in DAF-3 or DAF-16 expression.

Alternatively, the effect of candidate modulators on expression or activity may be measured at the level of DAF-3 or DAF-16 protein production

using the same general approach in combination with standard immunological detection techniques, such as Western blotting or immunoprecipitation with a DAF-3 or DAF-16-specific antibody (for example, the DAF-3 or DAF-16 antibodies described herein). Again, useful anti-diabetic or anti-obesity therapeutic modulators are identified as those which produce a decrease in DAF-3 or DAF-16 polypeptide production. Antagonists may also affect DAF-3 or DAF-16 activity without any effect on expression level. For example, the identification of kinase cascades upstream of DAF-3 and DAF-16 (as described herein) suggest that the phosphorylation state of these polypeptides is correlated with activity. Phosphorylation state may be monitored by standard Western blotting using antibodies specific for phosphorylated amino acids. In addition, because DAF-3 and DAF-16 are transcription factors, reporter genes bearing operably linked DAF-3 or DAF-16 binding sites (for example, the myoII enhancer element) may be used to directly monitor the effects of antagonists on DAF-3 or DAF-16 gene activity.

Candidate modulators may be purified (or substantially purified) molecules or may be one component of a mixture of compounds (e.g., an extract or supernatant obtained from cells). In a mixed compound assay, DAF-3 or DAF-16 expression is tested against progressively smaller subsets of the candidate compound pool (e.g., produced by standard purification techniques, e.g., HPLC or FPLC; Ausubel et al., *supra*) until a single compound or minimal compound mixture is demonstrated to modulate DAF-3 or DAF-16 expression.

Candidate DAF-3 or DAF-16 antagonists include peptide as well as non-peptide molecules (e.g., peptide or non-peptide molecules found, e.g., in a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured).

Antagonists found to be effective at the level of cellular DAF-3 or DAF-16 expression or activity may be confirmed as useful in animal models (for

example, nematodes or mice). For example, the compound may ameliorate the glucose intolerance and diabetic symptoms of mouse models for Type II diabetes (e.g., a db mouse model), mouse models for Type I diabetes, or models of streptozocin- induced β cell destruction.

A molecule which promotes a decrease in DAF-3 or DAF-16 expression or DAF-3 or DAF-16 activity is considered particularly useful in the invention; such a molecule may be used, for example, as a therapeutic to decrease the level or activity of native, cellular DAF-3 or DAF-16 and thereby treat a glucose intolerance condition in an animal (for example, a human).

If desired, treatment with an antagonist of the invention may be combined with any other anti-diabetic or anti-obesity therapies.

Agonists

Also as discussed above, useful therapeutic compounds are those which up regulate the expression or activity of DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT. To isolate such compounds, expression of these genes is measured following the addition of candidate agonist molecules to a culture medium of DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT-expressing cells. Alternatively, the candidate agonists may be directly administered to animals (for example, nematodes or mice) and used to screen for effects on DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression.

DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT-expression is measured, for example, by standard Northern blot analysis (Ausubel et al., *supra*) using all or a portion of one of these genes as a hybridization probe. The level of DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression in the presence of the candidate molecule is compared to the level measured for the same cells, in the

same culture medium, or in a parallel set of test animals, but in the absence of the candidate molecule. Preferred modulators for anti-diabetic or anti-obesity purposes are those which cause an increase in DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression.

Alternatively, the effect of candidate modulators on expression may be measured at the level of DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT protein production using the same general approach in combination with standard immunological detection techniques, such as Western blotting or immunoprecipitation with an appropriate antibody. Again, the phosphorylation state of these polypeptides is indicative of DAF activity and may be measured on Western blots. Useful anti-diabetic or anti-obesity modulators are identified as those which produce an increase in DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT polypeptide production.

Candidate modulators may be purified (or substantially purified) molecules or may be one component of a mixture of compounds (e.g., an extract or supernatant obtained from cells). In a mixed compound assay, DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression is tested against progressively smaller subsets of the candidate compound pool (e.g., produced by standard purification techniques, e.g., HPLC or FPLC; Ausubel et al., *supra*) until a single compound or minimal compound mixture is demonstrated to modulate DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression.

Alternatively, or in addition, candidate compounds may be screened for those which agonize native or recombinant DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT activities. In one particular example, DAF-1 and DAF-4 phosphorylation of DAF-8 and DAF-14, or AKT phosphorylation of DAF-16, may be activated by agonists.

Candidate DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT agonists include peptide as well as non-peptide molecules (e.g., peptide or non-peptide molecules found, e.g., in a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured).

Agonists found to be effective at the level of cellular DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression or activity may be confirmed as useful in animal models (for example, nematodes or mice).

A molecule which promotes an increase in DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT expression or DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT activities is considered particularly useful in the invention; such a molecule may be used, for example, as a therapeutic to increase the level or activity of these native, cellular genes and thereby treat a glucose intolerance condition.

If desired, treatment with an DAF-1, DAF-2, DAF-4, DAF-7, DAF-8, DAF-11, DAF-14, AGE-1, or AKT agonist may be combined with any other anti-diabetic or anti-obesity therapies.

Animal Model Systems

Compounds identified as having activity in any of the above-described assays are subsequently screened in any number of available diabetic or obesity animal model systems, including, but not limited to ob (Coleman, *Dibetologia* 14: 141-148, 1978; Chua et al., *Science* 271: 994-996, 1996; Vaisse et al., *Nature Genet.* 14:95-100, 1996), db (Chen et al., Cell 84: 491-495, 1996), agouti mice, or fatty rats (Takaga et al. *Biochem. Biophys. Res. Comm.* 225: 75-83, 1996). Test compounds are administered to these animals according to standard methods. Additionally, test compounds may be tested in mice bearing

knockout mutations in the insulin receptor, IGF-1 receptor (e.g., Liu et al., Cell 75:59-72, 1993), IR-related receptor, DAF-7 homolog, or any of the *daf* (FKHR, AFX) genes described herein. Compounds can also be tested using any standard mouse or rat model of Type I diabetes, e.g., a streptozin ablated pancreas model.

In one particular example, the invention involves the administration of DAF-7 or its homolog as a method for treating diabetes or obesity. Evaluation of the effectiveness of such a compound is accomplished using any standard animal model, for example, the animal diabetic model systems described above. Because these mouse diabetic models are also associated with obesity, they provide preferred models for human obesity associated Type II diabetes as well. Such diabetic or obese mice are administered *C. elegans* or human DAF-7 according to standard methods well known in the art. Treated and untreated controls are then monitored for the ability of the compound to ameliorate the symptoms of the disease, e.g., by monitoring blood glucose, ketoacidosis, and atherosclerosis. Normalization of blood glucose and insulin levels is taken as an indication that the compound is effective at treating a glucose intolerance condition.

Therapy

Compounds of the invention, including but not limited to, DAF-7 and its homologs, and any antagonist or agonist therapeutic agent identified using any of the methods disclosed herein, may be administered with a pharmaceutically-acceptable diluent, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable formulations or compositions to administer such compositions to patients. Although intravenous administration is preferred, any appropriate route of administration may be employed, for example, parenteral, subcutaneous, intramuscular,

intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral administration. Therapeutic formulations may be in the form of liquid solutions or suspensions; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

Methods well known in the art for making formulations are found in, for example, "Remington's Pharmaceutical Sciences." Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for antagonists or agonists of the invention include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

DAF polypeptides are administered at any appropriate concentration, for example, for DAF-7, at a concentration of around 10nM.

Pedigree Analysis and Genetic Testing

The discovery described herein that DAF polypeptides are involved in glucose metabolism enables assays for genetic testing to identify those individuals with predispositions toward the development of glucose intolerance conditions, such as diabetes or obesity, by determining the presence of a mutation found in a human gene having identity to any of the *C. elegans daf-1*,

daf-2, *daf-3*, *daf-4*, *daf-7*, *daf-8*, *daf-11*, *daf-14*, *daf-16*, *age-1*, or *akt* genes. In one embodiment, the development of this testing method requires that the individual be a member of a family that has multiple affected and unaffected members carrying one mutation from the list of above-listed genes. Those skilled in the art will understand that a diabetic or obesity phenotype may be produced by *daf*, *age*, or *akt* mutations found on different chromosomes, and that low resolution genetic mapping of the diabetic condition in single family pedigrees will be sufficient to favor some *daf*, *age*, or *akt* genes over others as causing the glucose intolerance condition in a particular pedigree. In one particular example, mutations associated with glucose intolerance may be found in different genes in, for example, the DAF-7 signaling pathway in each pedigree. In addition, because mutations in a common pathway can show complex genetic interactions, multiple DAF mutations may segregate in single pedigrees. These mutations can behave recessively in some genetic backgrounds and dominantly in others.

Those skilled in the art further understand that, to determine disease linkage with a chromosomal marker, it may be necessary to evaluate the association of inheritance patterns of several different chromosomal markers (for example, from the collection of highly polymorphic mapped DNA allelic variants) in the genomic DNAs of family members and of the clinically affected individuals. Methods commonly used in determining segregation patterns of human genetic diseases are well known in the art. In addition, methods are known in the art for determining whether individuals in a family are useful for providing information to determine co-segregation of an allele with a glucose intolerance trait.

Here, fragments of genomic DNA (e.g., RFLP fragments) are prepared from each of the available members of the family, and each distinctive DNA allelic variant of the polymorphic chromosome marker within the family is

evaluated to determine which polymorphisms (i.e., chromosomal region) is linked with the glucose intolerance phenotype within a particular family. It is preferred that the parents of the marker individual be heterozyous for a DNA allelic variant so that the segregation pattern of the DNA allelic variant linked with the diabetic or obese phenotype in the marker can be recognized. The inheritance of the diabetic phenotype can be judged to be dominant or recessive, depending on the pattern of inheritance. Most diabetes is dominantly inherited, and therefore inbred pedigrees are generally not necessary in the etiology of the diabetic condition.

With respect to Type II diabetes, the highest rate of this kind of diabetes in the world is found in American Indians of the Pima tribe. Such families are useful for mapping one particular cause of diabetes, but, in general, diabetes is caused by mutations in a variety of genes, including *daf* genes. Thus, by testing for low resolution linkage to a candidate *daf*, *age*, or *akt* mutation, and then by sequencing the particular linked *daf* gene in affected and unaffected individuals, a particular *daf* mutation can be associated with a particular diabetic pedigree.

Human DAF homologues are mapped to chromosome regions using standard methods. For example, the probable DAF-16 homologue FKHR is located on chromosome 13, and AFX is located on the X chromosome. Any *daf*, *akt*, or *age* genes mapping to the approximate chromosomal regions associated with diabetes or glucose intolerance are sequenced from affected and unaffected individuals. Preferably, at least two genes per pedigree of 5-20 affected (and unaffected controls) are sequenced. The *daf* genomic regions are PCR amplified and compared between affected and unaffected DNA samples. Mutations detected in affected individuals are expected to (but need not) map to conserved domains of the DAF genes. Because it is known that not all carriers of known diabetes-inducing mutations show metabolic defects, we expect that

some non-diabetic non-glucose intolerant family members will carry the same *daf* mutation as affected family members. For this reason, a correlation of affected family members with a *daf* mutation is more important than a correlation of nonaffected with no mutation. Those skilled in the art will know that phenotypic classification of affected and unaffected individuals can greatly enhance the power of this genetic analysis (Nature Genet. 11: 241-247, 1995). In addition, other mutations in the same *daf* gene are expected in some but not all diabetic pedigrees. For dominant diabetic inheritance, the affected individuals carry a *daf*, *age*, or *akt* mutation as well as a normal allele. For recessive diabetic inheritance, individuals carry two *daf* mutations that may be identical or two independent mutations in the same gene. In addition, some diabetic individuals may carry mutations in more than one *daf*, *age*, or *akt* gene (so called non-allelic non-complementation).

It is routine in the art of genetic counseling to determine risk factors given

the presence of a closely linked molecular genetic marker in the genomic DNA of the individual and when combined with the additional understanding provided by the pedigree of the individual in the family. For example, a risk factor may be calculated for an individual in an *age*, *akt*, or *daf* chromosome family in a manner similar to those described for assessing the risk of other commonly known genetic diseases that are known to run in families, e.g., Huntington's disease and cystic fibrosis.

Once mutations in *daf*, *akt*, or *age* genes are associated with diabetes in a pedigree analysis, diagnostic PCR sequencing of these *daf* genes can be used to diagnose glucose intolerant, prediabetic, diabetic, obesity, and atherosclerotic conditions. Preferably, the *daf, akt, or age* gene regions are PCR amplified from patients and mutations detected in the *daf* genes using standard DNA sequencing or oligonucleotide hybridization techniques. The

use of such gene sequences or specific antibody probes to the products of these sequences provide valuable diagnostics, particularly in view of the likelihood there exist two classes of type II diabetics: those with defects in the TGF-β signaling genes, and those with defects in insulin signaling genes. Such genetic tests will influence whether drugs that affect DAF-7 TGF-β or DAF-2 insulin like signals are prescribed.

To carry out the above analysis (as well as the other screening, diagnostic, and therapeutic methods described herein), mammalian homologs corresponding to the *C. elegans daf-1*, *age-1*, *daf-4*, *daf-8*, and *daf-7* genes are isolated as described above for *daf-2*, *daf-3*, and *daf-16*. Again, standard hybridization or PCR cloning strategies are employed, preferably utilizing conserved DAF, AGE, or AKT motifs for probe design followed by comparison of less conserved sequences flanking these motifs. Exemplary motifs for these genes are as follows:

DAF-1 (139 amino acid motif) (SEQ ID NO: 13)
274 TSGSGMGPTTLHKLTIGGQIRLTGRVGSGRFGNVSRGDYRGEAVA
VKVFNALDEPAFHKETEIFETRMLRHPNVLRYIGSDRVDTGFVTELWL
VTEYHPSGSLHDFLLENTVNIETYYNLMRSTASGLAFLHNQIGGSK 412

DAF-1 (62 amino acid motif) (SEQ ID NO: 14)
450 EDAASDIIANENYKCGTVRYLAPEILNSTMQFTVFESYQCADVYSF
SLVMWETLCRCEDGDV 511

DAF-1 (31 amino acid motif) (SEQ ID NO: 15)
416 KPAMAHRDIKSKNIMVKNDLTCAIGDLGLSL 466

DAF-1 (72 amino acid motif) (SEQ ID NO: 16)

520 IPYIEWTDRDPQDAQMFDVVCTRRLRPTENPLWKDHPEMKHIMEII
KTCWNGNPSARFTS YICRKRMDERQQ 591

AGE-1 (150 amino acid motif) (SEQ ID NO: 17)
991 YFESVDRFLYSCVGYSVATYIMGIKDRHSDNLMLTEDGKYVHIDF
GHILGHGKTKLGIQRDRQPFILTEHFMTVIRSGKSVDGNSHELQKFKTL
CVEAYEVMWNNRDLFVSLFTLMLGMELPELSTKADLDHLKKTLFCNG
ESKEEARKF 1140

AGE-1 (113 amino acid motif) (SEQ ID NO: 18)
826 SPLDPVYKLGEMIIDKAIVLGSAKRPLMLHWKNKNPKSDLHLPFCA
MIFKNGDDLRQDMLVLQVLEVMDNIWKAANIDCCLNPYAVLPMGEMI
GIIEVVPNCKTIFEIQVGTG 938

AGE-1 (106 amino acid motif) (SEQ ID NO: 19)
642 LAFVWTDRENFSELYVMLEKWKPPSVAAALTLLGKRCTDRVIRKF
AVEKLNEQLSPVTFHLFILPLIQALKYEPRAQSEVGMMLLTRALCDYRI
GHRLFWLLRAEI 747

AGE-1 (60 amino acid motif) (SEQ ID NO: 38)
91 EIKLSDFKHQLFELIAPMKWGTYSVKPQDYVFRQLNNFGEIEVIFND
DQPLSKLELHGTF 150

AKT (121 amino acid motif) (SEQ ID NO: 60)
33685 QVLDDHDYGRCVDWWGVGVVMYEMMCGRLPFYSKDHNKLF
ELIMAGDLRFPSKLSQEARTLLTGLLVKDPTQRLGGGPEDALEICRADF

FRTVDWEATYRKEIEPPYKPNVQSETDTSYFD 34047

AKT (66 amino acid motif) (SEQ ID NO: 61)

32314 TMEDFDFLKVLGKGTFGKVILCKEKRTQKLYAIKILKKDVIIARE

EVAHTLTENRVLQRCKHPFLT 32511

AKT (45 amino acid motif) (SEQ ID NO: 62)

33509 KLENLLLDKDGHIKIADFGLCKEEISFGDKTSTFCGTPEYL

APEV 33643

AKT (57 amino acid motif) (SEQ ID NO: 63)

32667

YFQELKYSFQEQHYLCFVMQFANGGELFTHVRKCGTFSEPRARFY

GAEIVLALGYLH 32837

AKT (59 amino acid motif) (SEQ ID NO: 64)

31846

STFAIFYFQTMLFEKPRPNMFMVRCLQWTTVIERTFYAESAEVRQ

RWIHAIESISKKYK 32022

AKT (33 amino acid motif) (SEQ ID NO: 65)

33156 LQELKYSFQTNDRLCFVMEFAIGGDLYYHLNRE 33254

AKT (21 amino acid motif) (SEQ ID NO: 66)

30836 VVIEGWLHKKGEHIRNWRPRF 30898

AKT (26 amino acid motif) (SEQ ID NO: 67)

33276 FSEPRARFYGSEIVLALGYLHANSIV 33353

DAF-4 (139 amino acid motif) (SEQ ID NO: 20)

380 EYWIVTEFHERLSLYELLKNNVISITSANRIIMSMIDGLQFLHDDRPY
FFGHPKKPIIHRDIKSKNILVKSDMTTCIADFGLARIYSYDIEQSDLLGQV
GTKRYMSPEMLEGATEFTPTAFKAMDVYSMGLVMWEVISR 518

DAF-4 (61 amino acid motif) (SEQ ID NO: 21)

537 IGFDPTIGRMRNYVVSKKERPQWRDEIIKHEYMSLLKKVTEEMWD
PEACARITAGCAFARV 597

DAF-4 (20 amino acid motif) (SEQ ID NO: 22)

305 PITDFQLISKGRFGKVFKAQ 324

DAF-8 (163 amino acid motif) (SEQ ID NO: 23)

382 TDSETRSRFSLGWYNNPNRSPQTAEVRGLIGKGVRFYLLAGEVYVE
NLCNIPVFVQSIGANMKNGFQLNTVSKLPPTGTMKVFDMRLFSKQLRT
AAEKTYQDVYCLSRMCTVRVSFCKGWGEHYRRSTVLRSPVWFQAHL
NNPMHWVDSVLTCMGAPPRICSS 544

DAF-8 (44 amino acid motif) (SEQ ID NO: 24)

91 RAFRFPVIRYESQVKSILTCRHAFNSHSRNVCLNPYHYRWVELP 134

DAF-8 (38 amino acid motif) (SEQ ID NO: 25)

341 VEYEESPSWLKLIYYEEGTMIGEKADVEGHHCLIDGFT 378

DAF-14 (39 amino acid motif) (SEQ ID NO: 68)

9709 IRVSFCKGFGETYSRLKVVNLPCWIEIILHEPADEYDTV 9825

DAF-14 (45 amino acid motif) (SEQ ID NO: 69)

9409 SRNSKSSQIRNTVGAGIQLAYENGELWLTVLTDQIVFVQCPFLNQ
9543

DAF-14 (29 amino acid motif) (SEQ ID NO: 70)
9160 NEMLDPEPKYPKEEKPWCTIFYYELTVRV 9246

DAF-14 (29 amino acid motif) (SEQ ID NO: 71)
9307 QLGKAFEAKVPTITIDGATGASDECRMSL 9393

DAF-12 (105 amino acid motif) (SEQ ID NO: 72)
103 SPDDGLLDSSEESRRRQKTCRVCGDHATGYNFNVITCESCKAFFRR
NALRPKEFKCPYSEDCEINSVSRRFCQKCRLRKCFTVGMKKEWILNEEQ
LRRRKNSRLN 207

DAF-12 (89 amino acid motif) (SEQ ID NO: 73)
109
LDSSEESRRRQKTCRVCGDHATGYNFNVITCESCKAFFRRNALRPKE
FKCPYSEDCEINSVSRRFCQKCRLRKCFTVGMKKEWILNEEQ 197

DAF-12 (73 amino acid motif) (SEQ ID NO: 74)
551 DIMNIMDVTMRRFVKVAKGVPAFREVSQEGKFSLLKGGMIEMLTV
RGVTRYDASTNSFKTPTIKGQNVSVNVD 623

DAF-11 (112 amino acid motif) (SEQ ID NO: 75)
708 SGSLVDLMIKNLTAYTQGLNETVKNRTAELEKEQEKGDQLLMELL
PKSVANDLKNGIAVDPKVYENATILYSDIVGFTSLCSQSQPMEVVTLLS
GMYQRFDLIISQQGGYKV 819

DAF-11 (107 amino acid motif) (SEQ ID NO: 76)

825 METIGDAYCVAAGLPVVMEKDHVKSICMIALLQRDCLHHFEIPHR PGTFLNCRWGFNSGPVFAGVIGQKAPRYACFGEAVILASKMESSGVED RIQMTLASQQLLEE 931

DAF-11 (43 amino acid motif) (SEQ ID NO: 77)

520 DILKGLEYIHASAIDFHGNLTLHNCMLDSHWIVKLSGFGVNRL 562

DAF-11 (15 amino acid motif) (SEQ ID NO: 78)

618 DMYSFGVILHEIILK 632

DAF-7 (60 amino acid motif) (SEQ ID NO: 26)

290 NLAETGHSKIMRAAHKVSNPEIGYCCHPTEYDYIKLIYVNRDGRVS IANVNGMIAKKCGC 349

DAF-7 (20 amino acid motif) (SEQ ID NO: 27)

265 DWIVAPPRYNAYMCRGDCHY 284

DAF-7 (43 amino acid motif) (SEQ ID NO: 28)

240 VCNAEAQSKGCCLYDLEIEFEKIGWDWIVAPPRYNAYMCRGDC 282

DAF-7 (70 amino acid motif) (SEQ ID NO: 29)

281 DCHYNAHHFNLAETGHSKIMRAAHKVSNPEIGYCCHPTEYDYIKLI YVNRDGRVSIANVN GMIAKKCGCS 350

DAF-7 (35 amino acid motif) (SEQ ID NO: 30)

250 CCLYDLEIEFEKIGWDWIVAPPRYNAYMCRGDCHY 284

DAF-7 (13 amino acid motif)(SEQ ID NO: 51)
GWDWIVAPPRYNA

In one particular example, mammalian DAF-7 may be identified using the sub-domain amino acids 314-323. Exemplary degenerate oligonucleotides designed to PCR amplify this domain or hybridize (for example, as described in Burglin et al., (Nature 341:239-243, 1989) are as follows:

aa 263 oligo: GGNTGGGAYTRNRTNRTNGCNCC (23-mer, 16,000-fold degeneracy) (SEQ ID NO: 31)

aa 314 oligo: TGYTGYNNNCCNACNGAR (18-mer, 8000-fold degeneracy) (SEQ ID NO: 32).

The DNA sequence between the oligonucleotide probes is determined, and those sequences having the highest degree of homology are selected. Once isolated, these sequences are then tested in a *C. elegans daf-7* mutant or mouse model as described above for the ability to functionally complement the mutation or ameliorate the glucose intolerance phenotype.

Other Embodiments

In other embodiments, the invention includes any protein which possesses the requisite level of amino acid sequence identity (as defined herein) to DAF-2, DAF-3, or a DAF-16 sequence; such homologs include other substantially pure naturally-occurring mammalian DAF polypeptides (for example, human DAF polypeptides) as well as allelic variants; natural mutants; induced mutants; proteins encoded by DNA that hybridizes to the DAF DNA

sequence or degenerate conserved domains of DAF proteins (e.g., those described herein) under high stringency conditions; and proteins specifically bound by antisera directed to a DAF-2, DAF-3, or DAF-16 polypeptide.

The invention further includes analogs of any naturally-occurring DAF-2, DAF-3, or DAF-16 polypeptides. Analogs can differ from the naturally-occurring protein by amino acid sequence differences which do not destroy function, by post-translational modifications, or by both. Modifications include *in vivo* and *in vitro* chemical derivatization of polypeptides, e.g., acetylation, carboxylation, phosphorylation, or glycosylation; such modifications may occur during polypeptide synthesis or processing or following treatment with isolated modifying enzymes. Analogs can also differ from the naturally-occurring DAF polypeptide by alterations in primary sequence. These include genetic variants, both natural and induced (for example, resulting from random mutagenesis by irradiation or exposure to ethanemethylsulfate or by site-specific mutagenesis as described in Sambrook, Fritsch and Maniatis, Molecular Cloning: A Laboratory Manual (2d ed.), CSH Press, 1989, or Ausubel et al., *supra*). Also included are cyclized peptides, molecules, and analogs which contain residues other than L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids.

In addition to full-length polypeptides, the invention also includes DAF-2, DAF-3, and DAF-16 polypeptide fragments. As used herein, the term "fragment," means at least 20 contiguous amino acids, preferably at least 30 contiguous amino acids, more preferably at least 50 contiguous amino acids, and most preferably at least 60 to 80 or more contiguous amino acids. Fragments of such DAF polypeptides can be generated by methods known to those skilled in the art or may result from normal protein processing (e.g., removal of amino acids from the nascent polypeptide that are not required for biological activity or removal of amino acids by alternative mRNA splicing or

alternative protein processing events).

For certain purposes, all or a portion of the DAF-2, DAF-3, or DAF-16 polypeptide sequence may be fused to another protein (for example, by recombinant means). In one example, the DAF polypeptide may be fused to the green fluorescent protein, GFP (Chalfie et al., *Science* **263**:802-805, 1994). Such a fusion protein is useful, for example, for monitoring the expression level of the DAF polypeptide *in vivo* (for example, by fluorescence microscopy) following treatment with candidate or known DAF agonists or antagonists.

The methods of the invention may be used to diagnose or treat any condition related to glucose intolerance or obesity in any mammal, for example, humans, domestic pets, or livestock. Where a non-human mammal is diagnosed or treated, the DAF polypeptide, nucleic acid, or antibody employed is preferably specific for that species.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

Other embodiments are within the following claims.

What is claimed is:

Claims

1. A method of screening for a compound that decreases the activity of a DAF polypeptide, said method comprising the steps of:

(a) exposing a non-human transgenic animal, whose cells comprise a transgene coding for a mammalian DAF polypeptide, to said compound; and

(b) determining the activity of said DAF polypeptide in said transgenic animal, a decrease in DAF polypeptide activity as compared to untreated controls being indicative of a compound that is capable of decreasing DAF polypeptide activity.

2. The method of claim 1, wherein said animal is a nematode.

3. The method of claim 1, wherein said nematode carries a mutation in the corresponding endogenous *daf* gene, said mutation decreasing or eliminating the activity of said endogenous nematode *daf* gene product.

4. A method for identifying a modulatory compound that is capable of decreasing the expression or activity of a *daf* gene, involving:

(a) providing a cell expressing said *daf* gene; and

(b) contacting said cell with a candidate compound, a decrease in *daf* expression or activity following contact with said candidate compound identifying a modulatory compound.

5. The method of claim 1 or 4, wherein said compound is capable of treating an impaired glucose tolerance condition or obesity.

6. The method of claim 1 or 4, wherein said compound is capable of decreasing the expression or activity of DAF-3 or DAF-16.

7. A method for the identification of a compound that is capable of ameliorating or delaying an impaired glucose tolerance condition or obesity, comprising the steps of:

(a) providing a dauer larvae comprising a mutation in a *daf* gene; and

(b) contacting said dauer larvae with a compound, wherein release from the dauer larval state is an indication that said compound is capable of ameliorating or delaying an impaired glucose tolerance condition or obesity.

8. A method for the identification of a compound for ameliorating or delaying an impaired glucose tolerance condition or obesity, said method comprising the steps of:

(a) providing a *daf-2*, *daf-16* mutant nematode;

(b) expressing in the cells of said nematode a mammalian DAF-16 polypeptide, whereby said nematode forms a dauer larva; and

(c) contacting said dauer larva with a compound, a release from the dauer larval state being an indication that said compound is capable of ameliorating or delaying said glucose intolerance condition or obesity.

9. A method of determining whether a human gene is involved in an impaired glucose tolerance condition or obesity, comprising the steps of :

(a) providing a nematode having a mutation in a *daf* or *age* gene; and

(b) expressing in said nematode said human gene operatively linked to a nematode gene promoter, wherein complementation of said *daf* or *age* mutation in said nematode is indicative of a human gene that is involved in an impaired glucose tolerance condition or obesity.

10. Isolated DNA encoding a DAF-16 polypeptide, said DNA complementing a DAF-16 mutation in *C. elegans*.

11. The DNA of claim 10, wherein said DNA encodes a polypeptide that complements an FKHR or AFX mutation in a mouse.

12. A method of detecting a gene or a portion thereof found in a human cell having sequence identity to the *daf-16* sequence of Figs. 13A or 13B, said method comprising:

(a) contacting DNA encoding a nematode DAF16 polypeptide or a portion thereof greater than about 12 residues in length, or a degenerate oligonucleotide corresponding to SEQ ID NOS: 54, 55, 56 or 57, with a preparation of DNA from said human cell under hybridization conditions providing detection of DNA sequences having about 70% or greater nucleic acid sequence identity to the *daf-16* sequence of Figs. 13A or 13B; and

(b) isolating said human gene or portion thereof.

13. A method of isolating a gene or a portion thereof found in a human cell having 90% nucleic acid sequence identity to a sequence encoding SEQ ID NOS: 54, 55, 56, or 57, said method comprising:

(a) amplifying by PCR said human gene or portion thereof using oligonucleotide primers wherein said primers

(i) are each greater than about 12 residues in length;

(ii) each have regions of complementarity to opposite DNA strands in a region of the nucleotide sequence of Figs. 13A or 13B; and

(b) isolating said human gene or portion thereof.

14. The method of claim 12 or 13, wherein said method further comprises testing said gene or portion thereof for the ability to functionally complement a *C. elegans daf-16* mutant.

15. A method for ameliorating or delaying the onset of an impaired glucose tolerance condition or obesity in a patient, said method comprising administering to said patient a therapeutically effective amount of a compound that is capable of inhibiting the activity of a DAF-16 or DAF-3 polypeptide.

16. A method for ameliorating or delaying the onset of an impaired glucose tolerance condition or obesity in a patient, said method comprising administering to said patient a therapeutically effective amount of a DAF polypeptide.

17. The method of claim 16, wherein said DAF polypeptide is a nematode or human DAF-7 polypeptide.

18. A method for the identification of a compound for ameliorating or delaying an impaired glucose tolerance condition or obesity, said method comprising the steps of:

(a) providing a *daf-7*, *daf-3* mutant nematode;

(b) expressing in the cells of said nematode a mammalian DAF-3 polypeptide, whereby said nematode forms a dauer larva; and

(c) contacting said dauer larva with a compound, a release from the dauer larval state being an indication that said compound is capable of ameliorating or delaying said glucose intolerance condition or obesity.

19. A method of detecting a gene or a portion thereof found in a human cell having sequence identity to any of the *daf-3* sequences of Figs. 11A, 11B, or 11C, said method comprising:

(a) contacting DNA encoding a nematode DAF3 polypeptide or a portion thereof greater than about 12 residues in length, or a degenerate

oligonucleotide corresponding to SEQ ID NOS: 35, 36, or 85 with a preparation of DNA from said human cell under hybridization conditions providing detection of DNA sequences having about 70% or greater nucleic acid sequence identity to any of the *daf-3* sequences of Figs. 11A, 11B, or 11C; and

(b) isolating said human gene or portion thereof.

20. A method of isolating a gene or a portion thereof found in a human cell having 90% nucleic acid sequence identity to a sequence encoding SEQ ID NOS: 35, 36, or 85, said method comprising:

(a) amplifying by PCR said human gene or portion thereof using oligonucleotide primers wherein said primers

(i) are each greater than about 12 residues in length; and

(ii) each have regions of complementarity to opposite DNA strands in a region of any of the nucleotide sequences of Figs. 11A, 11B, or 11C; and

(b) isolating said human gene or portion thereof.

21. The method of claim 19 or 20, wherein said method further comprises a step of testing said gene or portion thereof for the ability to functionally complement a *C. elegans daf-3* mutant.

22. A method of producing a transgenic non-human animal, said animal lacking an endogenous *daf* gene and being capable of expressing a human DAF polypeptide, said method comprising the steps of:

(a) providing a transgenic non-human animal whose germ cells and somatic cells contain a mutation in a *daf* gene; and

(b) introducing a transgene encoding a human DAF polypeptide into an

embryonal cell of said non-human animal, said transgene being capable of expressing said human polypeptide.

23. A method of diagnosing an impaired glucose tolerance condition or obesity, or a propensity thereto, in a mammal, said method comprising analyzing the DNA of said mammal to determine whether said DNA contains a mutation in a DAF gene, wherein the identification of said mutation indicates that said mammal suffers from an impaired glucose tolerance condition or obesity, or a propensity thereto.

1/53

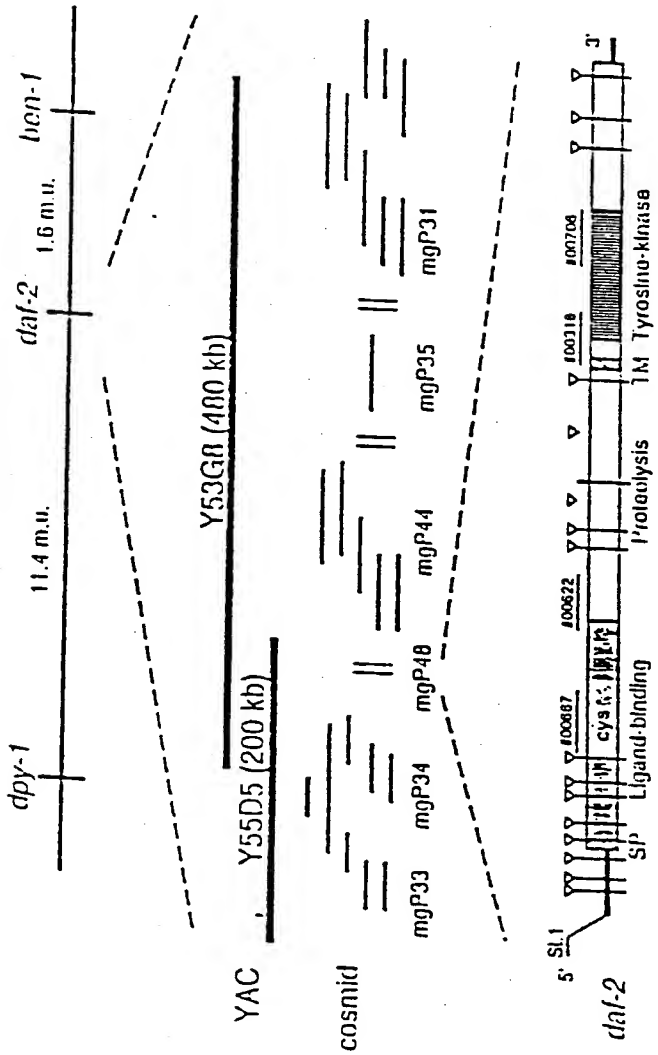


Fig.1

Fig. 2A

The *daf*-2 cDNA sequen

```
   1 ggttttaatta ctcaagtttg agctccaaga gcacacactt gatcgtcgga
  51 ttctactgta ctccccgaaa aaccaacaaa aaacacaagt ttttgaacac
 101 ttgtaaatgc agacagaacg atgacgagaa tgaatatttgt cagatgtcgg
 151 agacgacaca aaattttgga aaatttggaa gaagagaatc tcggcccgag
 201 ctgctcgtcg acgacttcaa caaccgctgc caccgaagct ctcggaacaa
 251 ccactgagga tatgaggctt aagcagcagc gaagctcgtc gcgtgccacg
 301 gagcacgata ttgtcgacgg caattaccac gacgacgagc acatcacaat
 351 gagacggctt cgacttgtca aaaattcgcg gacgcggcgt agaacgacgc
 401 ccgattcaag tatggactgc tatgaggaaa acccgccatc acaaaaactt
 451 caataaatta ttcttggatt tctaaaaagt catcaatgac gtcattaatg
 501 cttttactgc tattcgcttt tgtacagccg tgtgccttaa tagtcgaaaa
 551 acgatgcggc ccaatcgata ttcgaaatag gccgtgggat attaagccgc
 601 aatggtcgaa acttggtgat ccgaacgaaa aagatttggc tggtcagaga
 651 atggtcaact gcacagtggt ggaaggttcg ctgacaattt catttgtact
 701 gaaacacaag acaaaagcac aagaagaaat gcatcgaagt ctacagccaa
 751 gatattccca agacgaattt atcactttta cgcatctacg tgaaattact
 801 ggaacttctgc tcgtttttga gactgaagga ttagtggatt tgcgtaaaat
 851 tttcccaaat cttcgtgtaa ttggaggccg ttcgctgatt caacactatg
 901 cgctgataat ttatcgaaat ccggatttgg agatcggtct tgacaagctt
 951 tccgtaattc gaaatggtgg tgtacggata atcgataatc gaaaactgtg
1001 ctacacgaaa acgattgatt ggaaacattt gatcacttct tccatcaacg
1051 atgttgtcgt tgataatgct gccgagtacg ctgtcactga gactggattg
1101 atgtgcccac gtggagcttg cgaagaggat aaaggcgaat caaagtgtca
1151 ttattttggag gaaaagaatc aggaacaagg tgtcgaaaga gttcagagtt
1201 gttggtcgaa caccacttgc caaaagtctt gtgcttatga tcgtctttctt
1251 ccaacgaaag aaatcggacc gggatgtgat gcgaacggcg atcgatgtca
1301 cgatcaatgc gtgggcggtt gtgagcgtgt gaatgatgcc acagcatgcc
1351 acgcgtgcaa gaatgtctat cacaagggaa agtgtatcga aaagtgtgat
1401 gctcacctgt accttctcct tcaacgtcgt tgtgtgaccc gtgagcagtg
1451 tctgcagctg aatccggtgc tctcgaacaa aacagtgcct atcaaggcga
1501 cggcaggcct ttgctcggat aaatgtcccg atggttatca aattcaaccg
1551 gatgatcatc gagaatgccg aaaatgcgtt ggcaagtgtg agattgtgtg
1601 cgagatcaat cacgtcattg atacgtttcc gaaggcacag gcgatcaggc
1651 tatgcaatat tattgacgga aatctgacga tcgagattcg cggaaaacag
1701 gattcgggaa tggcgtccga gttgaaggat atatttgcga acattcacac
1751 gatcaccggc tacctgttgg tacgtcaatc gtcaccgttt atctcgttga
1801 acatgttccg gaatttacga cgtattgagg caaagtcact gttcagaaat
1851 ctatatgcta tcacagtttt tgaaaatccg aatttaaaaa agctattcga
1901 ttcaacgacg gatttgacgc ttgatcgtgg aactgtgtca attgccaata
1951 acaagatgtt atgcttcaag tatatcaagc agctaatgtc aaagttaaat
2001 ataccactcg atccgatага tcaatcagaa gggacaaatg gtgagaaggx
2051 aatctgtgag gatatggcaa tcaacgtgag catcacagcg gttcaacgcgg
2101 actcggtctt ctttagttgg ccctcattca acattaccga tatagatcag
2151 cgaaagtttc tcggctacga gctctttctc aaagaagtcc cacgaatcga
2201 tgagaacatg acgatcgaag aggatcgaag tgcgtgtgtc gattcgtggc
2251 agagtgtctt caaacagtac tacgagacgt cgaacggtga accgaccccg
2301 gacatttttа tggatattgg accgcgcgag cgaattcggc cgaatacgct
2351 ctacgcgtac tatgtggcga cgcagatggt gttgcatgcc ggtgcgaaga
2401 acggtgtatc gaagattggt tttgtgagga cgagctacta tacgcctgat
2451 cctccgacgt tggcactagc gcaagtcgat tcggacgcta ttcatattac
2501 gtgggaagcg ccgctccaac cgaacggaga cctcacgcat tacacaatta
2551 tgtggcgtga gaatgaagtg agcccgtacg aggaagccga aaagttttgt
2601 acagatgcaa gcacccccgc aaatcgacaa cgcacgaaag atccgaaaga
2651 gacgattgta gccgataagc cagtcgatat tccgtcatca cgtaccgtag
2701 ctccgacact tttgactatg atgggtcacg aagatcagca gaaaacgtgc
```

Fig. 2B (sheet 1 of 3)

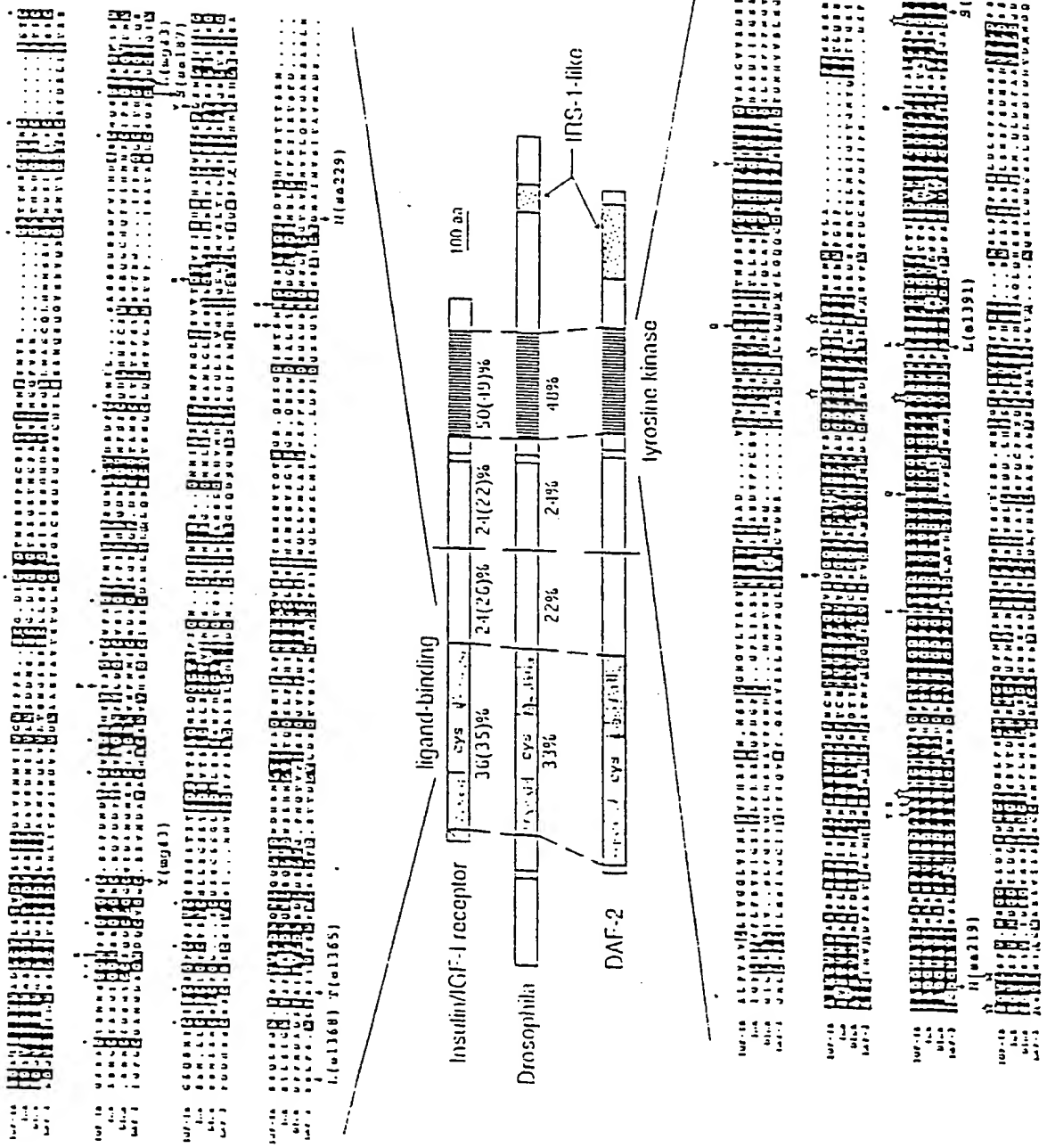
```
2751  gctgcaacgc  ;gttgttg ttcgtgttcg gctatcg  , aatcatcgga
2801  acagaacaag . gaagcgac cggatccgat gtcggcg .c gaatcatctg
2851  catttgagaa taagctgttg gatgaggttt taatgccgag agacacgatg
2901  cgagtgagac gatcaattga agacgcgaat cgagtcagtg aagagttgga
2951  aaaagctgaa aatttgggaa aagcttcaaa aacttctggt ggaaagaagc
3001  cgctgatcca tatttcgaag aagaagccgt cgagcagcag cacacatcc
3051  acaccggctc caacgatcgc atcaatgtat gccttaacaa ggaaaccgac
3101  tacggtgccg ggaacaagga ttcggcttcta cgagatctac gaacctttac
3151  ccggaagctg ggcgattaat gtatcagctc tggcattgga taatagttat
3201  gtgatacgaa atttgaagca ttacacactt tatgcgattt cttatccgc
3251  gtgccaaaac atgacagtac ccggagcatc ttgctcaata tcccatcgtg
3301  cgggagcatt gaaacgaaca aaacacatca cagacattga taaagtgttg
3351  aatgaaacaa ttgaatggag atttatgaat aatagtcaac aagtcaacgt
3401  gacgtgggat ccaccgactg aagtgaatgg tggaatattc ggttatgttg
3451  taaagcttaa gtcaaaagtc gatggatcaa ttgttatgac gagatgtgtc
3501  ggtgcgaaga gaggatattc aacacggaat cagggtgtcc tattccagaa
3551  tttggccgat ggacgttatt ttgtctcagt aacggcgacc tctgtacacg
3601  gcgctggacc ggaagccgaa tcctccgacc caatcgtcgt catgacgcca
3651  ggctttcttca ctgtggaaat cattctcggc atgcttctcg tcttttttgat
3701  tttaatgtca attgccggtt gtataatcta ctactacatt caagtacgct
3751  acggcaaaaa agtgaaagct ctatctgact ttatgcaatt gaatcccgaa
3801  tattgtgtgg acaataagta caatgcagac gattgggagc tacggcagga
3851  tgatgttgtg ctcggacaac agtgtggaga gggatcattc ggaaaagtgt
3901  acctaggaac tggaaataat gttgtttctc tgatgggtga tcgtttcgga
3951  ccgtgtgcta ttaagattaa tgtagatgat ccagcgtcga ctgagaatct
4001  caactatctc atggaagcta atattatgaa gaactttaag actaacttta
4051  tcgtccaact gtacggagtt atctctactg tacaaccagc gatggttgtg
4101  atggaaatga tggatcttgg aaatctccgt gactatctcc gatcgaaacg
4151  cgaagacgaa gtgttcaatg agacggactg caacttttc gacataatcc
4201  cgagggataa attccatgag tgggccgcac agatttgtga tggtatggcg
4251  tacctggagt cgctcaagtt ttgccatcga gatctcgccg cacgtaattg
4301  catgataaat cgggatgaga ctgtcaagat tggagatttc ggaatggctc
4351  gtgatctatt ctatcatgac tattataagc catcgggcaa gcgtatgatg
4401  cctgttcgat ggatgtcacc cgagtcgttg aaagacggaa agtttgactc
4451  gaaatctgat gtttggagct tcggagttgt tctctatgaa atggttacac
4501  tcggtgctca gccatatatt ggtttgagta atgatgaggt gttgaattat
4551  attggaatgg cccggaaggt tatcaagaag cccgaatgtt gtgaaaacta
4601  ttggtataag gtgatgaaaa tgtgctggag atactcacct cgggatcgtc
4651  cgacgttcct ccagctcgtt catcttctag cagctgaagc ttcaccagaa
4701  ttccgagatt tatcatttgt cctaaccgat aatcaaatga tccttgacga
4751  ttcagaagca ctggatcttg atgatattga tgatactgat atgaatgatc
4801  aggttgtcga ggtggcaccg gatgttgaga acgtcgaggt tcagagtgat
4851  tcggaacgtc ggaatacgga ttcaataccg ttgaaacagt ttaagacgat
4901  ccctccgatc aatgcgacga cgagtcattc gacaatatcg attgatgaga
4951  caccgatgaa agcgaagcag cgagaaggat cgctggatga ggagtacgca
5001  ttgatgaatc atagtggagg tccgagtgat gcggaagttc ggacgtatgc
5051  tggtgatgga gattatgtgg agagagatgt tcgagagaat gatgtgccaa
5101  cgcgacgaaa tactggtgca tcaacatcaa gttacacagg tggtggtcca
5151  tattgcctaa caaatcgtgg tggttcaaat gaacgaggag ccggttttcgg
5201  tgaagcagta cgattaactg atggtgttgg aagtggacat ttaaatgatg
5251  atgattatgt tgaaaaagag atatcatcca tggatacgcg ccggagcacg
5301  ggcgccttcga gcttttccta cggtgttcca cagacgaatt ggagtggaaa
5351  tcgtggtgcc acgtattata cgagtaaagc tcaacaggca gcaactgcag
5401  cagcagcagc agcagcagct ctccaacagc aacaaaatgg tggtcgaggc
5451  gatcgattaa ctcaactacc cggaactgga catttacaat cgacacgtgg
5501  tggacaagat ggagattata ttgaaactga accgaaaaat tatagaaata
```

Fig. 2B (sheet 2 of 3)

```
5551  atggatctcc  .cgcgaaac ggcaacagcc gtgacat    caacggacgt
5601  tcggcttttcg .cgaaaatga gcatctaatc gaggataatg agcatcatcc
5651  actttgtctga aacccccaaa aaatcccgcc tcttaaatta taaattatct
5701  cccacattat catatcttcta cacgaatatc ggatttttt tcagatttt
5751  tctgaaaaat tctgaataat tttaccccat ttttcaaatc tctgtatttt
5801  ttttgttat tacccc
```

Fig. 2B (Sheet 3 of 3)

Fig. 2C


ligand-binding
tyrosine kinase
Insulin/IGF-1 receptor
Drosophila
DAF-2
IRS-1-like
100 aa
36(35)%
33%
24(26)%
22%
24(22)%
24%
50(49)%
40%

7/53

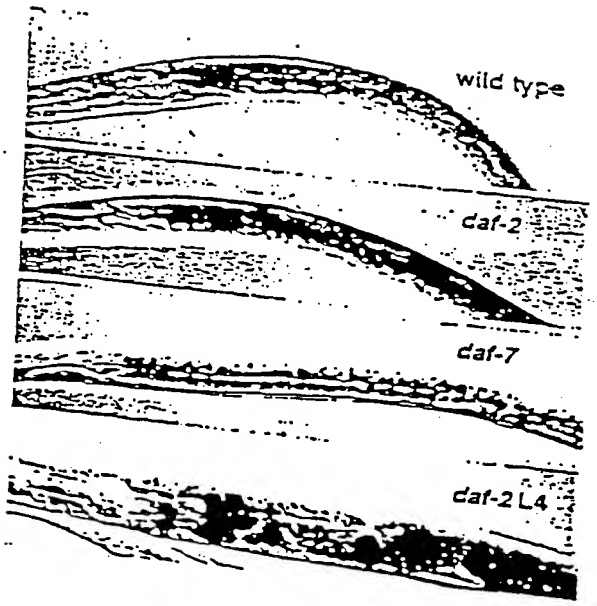


Fig. 3

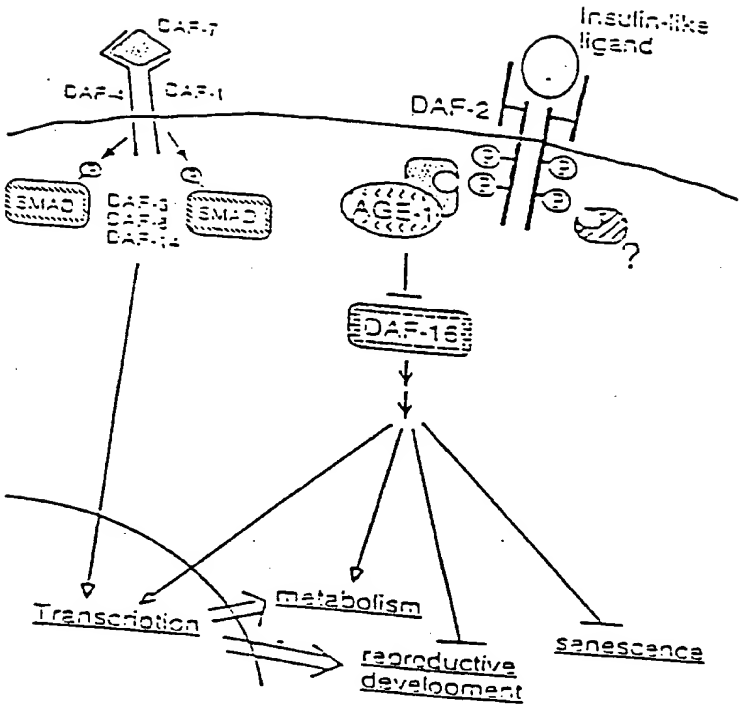

DAF-7
DAF-4
DAF-1
Insulin-like ligand
DAF-2
SMAD
DAF-3
DAF-8
DAF-14
SMAD
AGE-1
?
DAF-16
Transcription
metabolism
reproductive development
senescence

Fig. 4

FIG. 5A

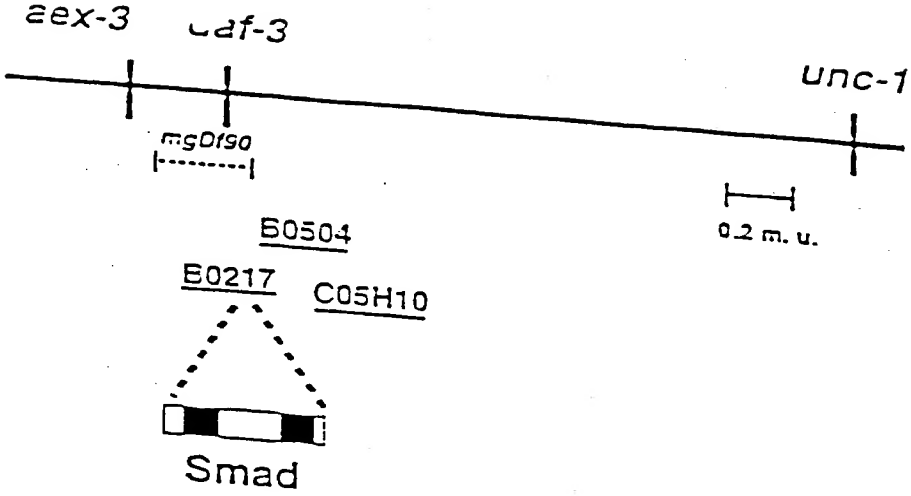


FIG. 5B

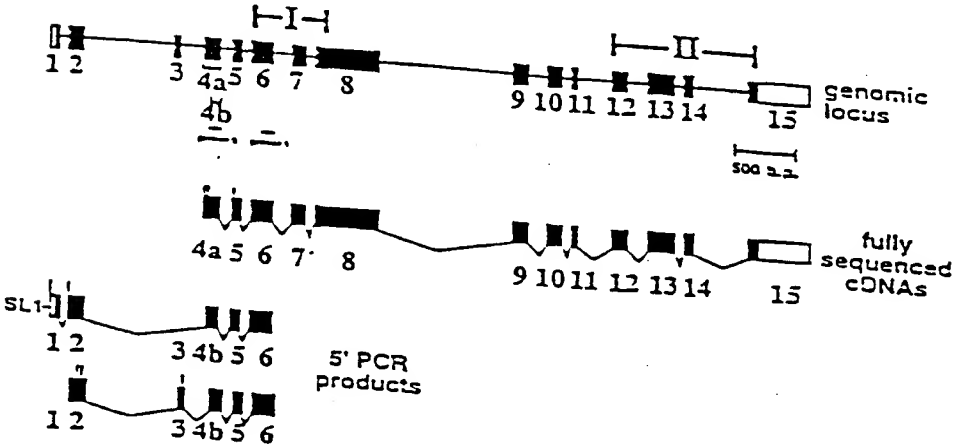


FIG. 5C

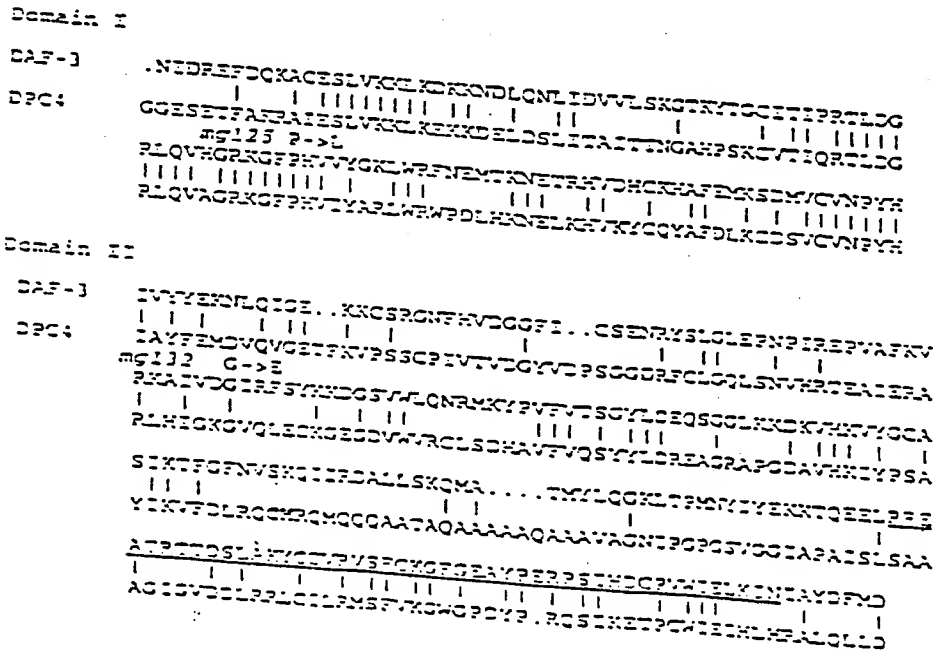


FIG. 6A

FIG. 6B

FIG. 6C

FIG. 6D

FIG. 6E

FIG. 6F

FIG. 6G

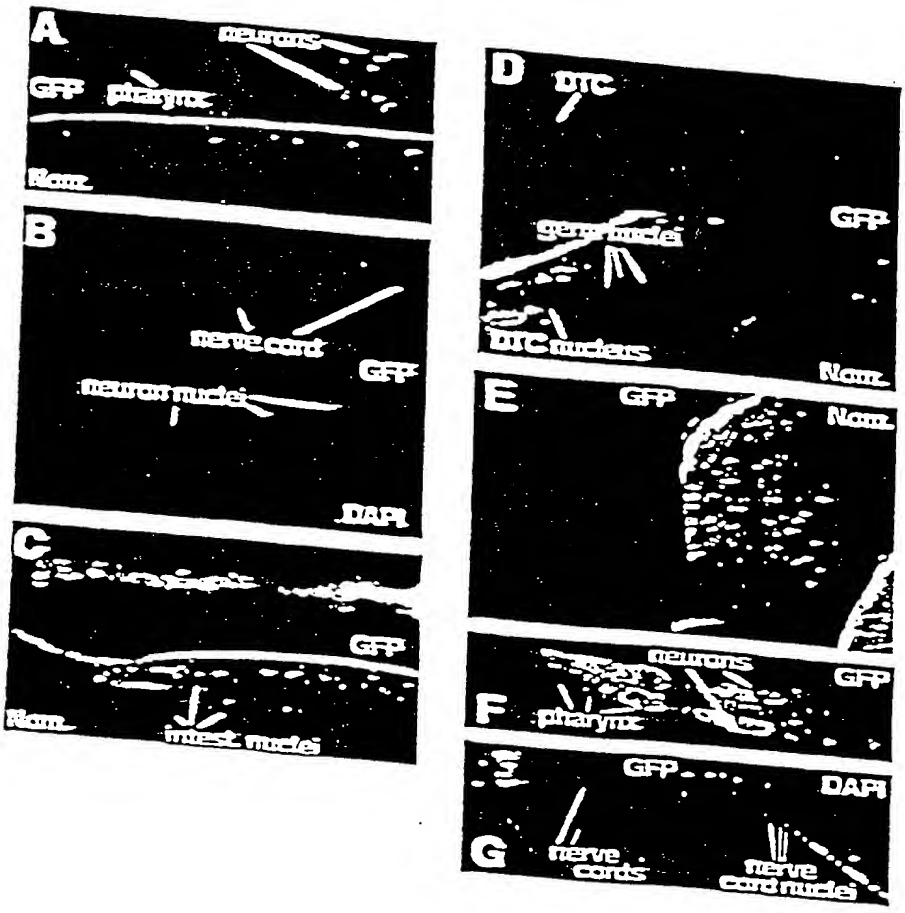


Fig. 6 A-G

	Rescue of *daf-7; daf-3* dauer (i.e. rescued)	Suppression of *daf-7* non-dauers (i.e. suppressed)
1	5 ± 1%	97 ± 3%
2	12 ± 3%	82 ± 6%
3	ND	90 ± 3%
4 Controls	0%	19 ± 4%

Fig. 7

FIG. 8

FIG. 8

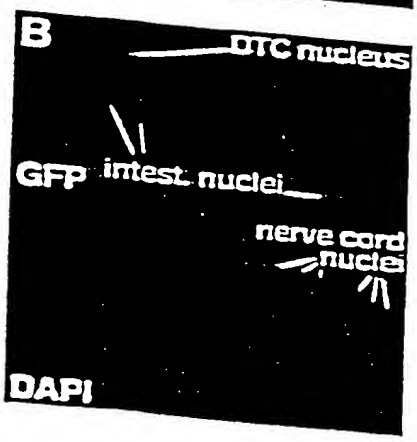


FIG. 8A-B

Fig. 9A

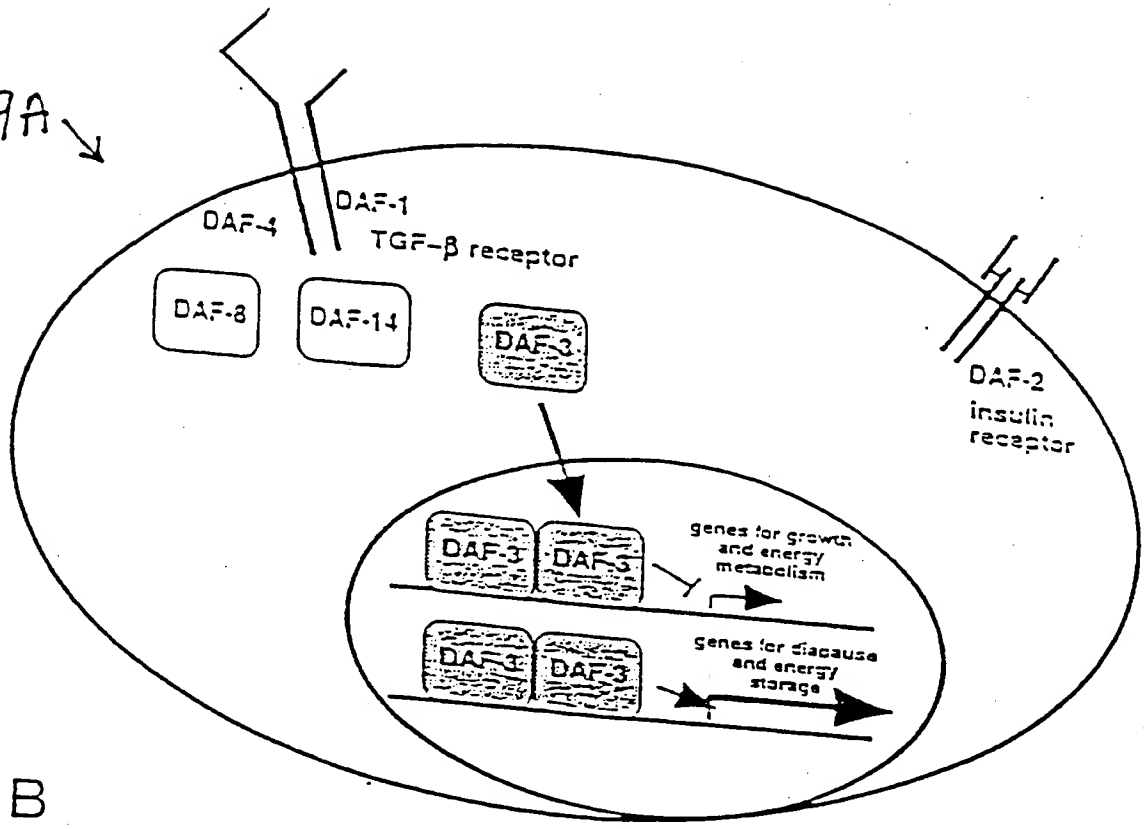


FIG. 9B

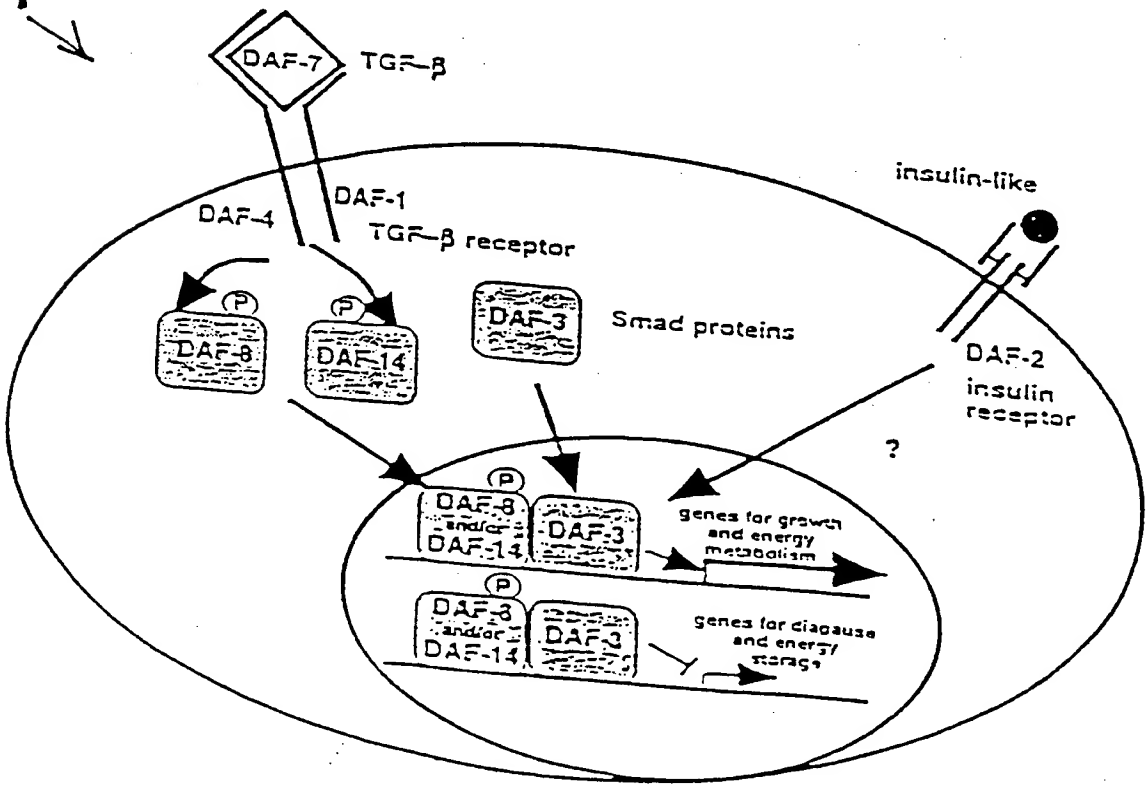
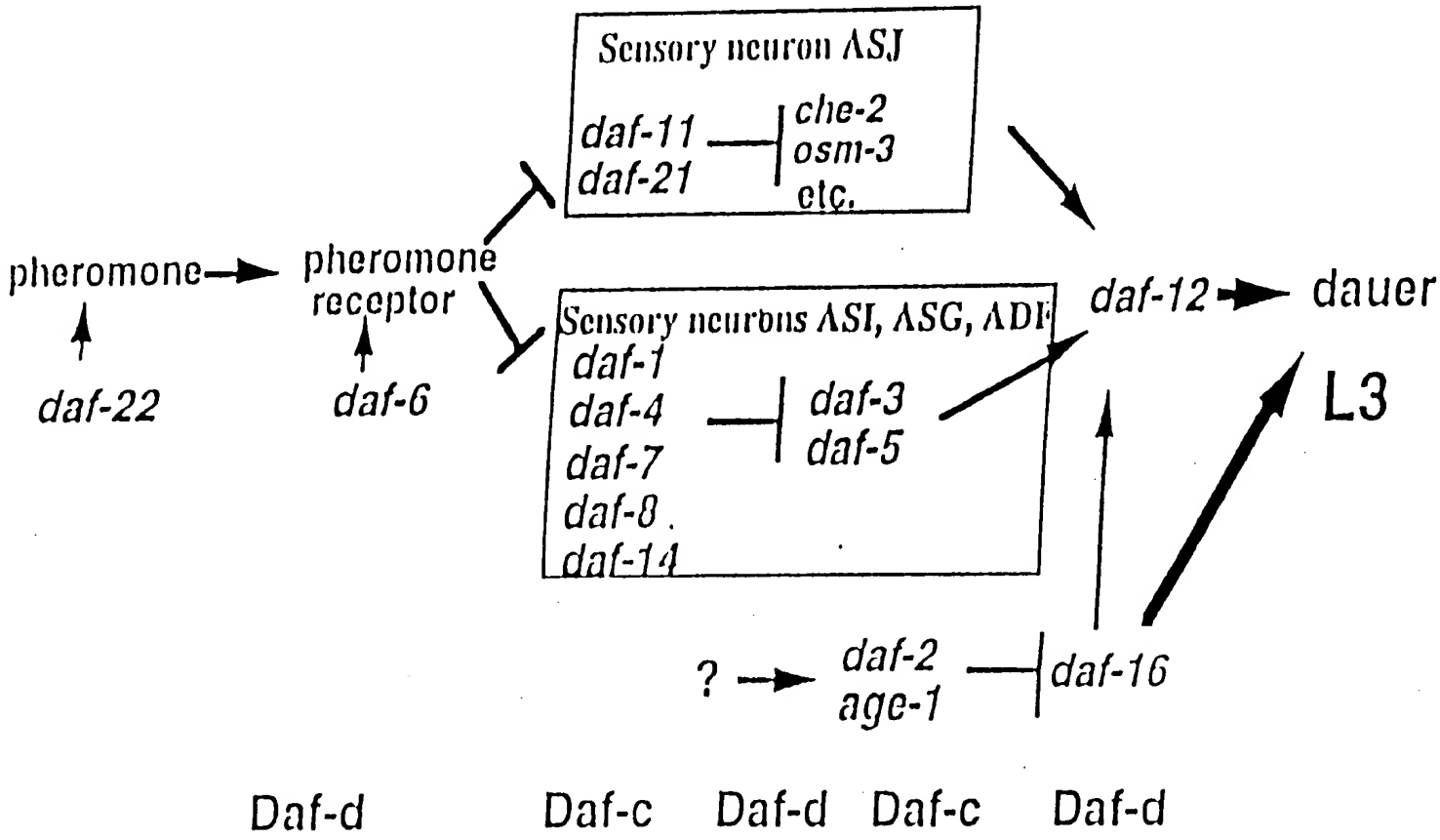


Fig. 10

The genetic pathway that regulates dauer formation



```
   1 atgaagctaa tagcaacttc tcttctagtt cccgacgagc acacaccgat
  51 gatgtcacca gtgaatacaa ctacaaagat tctacaacgg agtggtatta
 101 aaatggaaat cccgccatat ttggatccag acagtcagga tgatgacccg
 151 gaagatgggg tcaactaccc ggatccagat ttatttgaca caaaaaacac
 201 aaatatgacc gagtacgatt tggatgtgtt gaagcttgga aaaccagcag
 251 tagatgaagc acggaaaaag atcgaagttc ccgacgctag tgcgcccgcca
 301 aacaaaattg tagaatattt gatgtattat agaacgttaa aagaaagtga
 351 actcatacaa ctgaatgcgt atcggacaaa acgaaatcga ttatcgttga
 401 acttggtcaa aaacaattat gatcgagagt tcgaccaaaa agcttgcgag
 451 tcccctggtga aaaaattgaa ggataagaag aatgatctcc agaacctgat
 501 tgatgtggtt cttcaaaag gtacaaaata taccggttgc attacaattc
 551 caaggacact tgatggcccgg ttacaggtcc acggaagaaa aggttcccct
 601 cacgtagtct atggcaaact gtggaggttt aatgaaatga caaaaaacga
 651 aacgcgtcat gtggaccact gcaagcacgc atttgaaatg aaaagtgaca
 701 tggtatgcgt gaatccctat cactacgaaa ttgtcattgg aactatgatt
 751 gttgggcaga gggatcatga caatcgagat atgccgccgc cacatcaacg
 801 ctaccacact ccaggtcggc aggatccagt tgacgatatg agtagattta
 851 taccaccagc ttccattcgt ccgcctccga tgaacatgca cacaaggcct
 901 cagcctatgc ctcaacaatt gccttcagtt ggcgcaacgt ttgcccatcc
 951 tctcccacat caggcgccac ataacccagg ggtttcacat ccgtactcca
1001 ttgctccaca gacccattac ccgttgaaca tgaacccaat tccgcaaatg
1051 ccgcaaatgc cacaaatgcc accacctctc catcagggat atggaatgaa
1101 tgggccgagt tgctcttcag aaaacaacaa tccattccac caaaatcacc
1151 attataatga tattagccat ccaaatcact attcctacga ctgtggtccg
1201 aacttgtacg ggtttccaac tccttatccg gattttcacc atccttttcaa
1251 tcagcaacca caccagccgc cacaactatc acaaaaccat acgtcccaac
1301 aaggcagtca tcaaccaggg caccaaggtc aggtaccgaa tgatccacca
1351 attcaagac cagtgttaca accatcaaca gtcaccttgg acgtgttccg
1401 tcggtactgt agacagacat ttggaaatcg attttttgaa ggagaaagtg
1451 aacaatccgg cgcaattaat cggtctagta acaaattcat tgaagaattt
1501 gattcgccga tttgtggtgt gacagttgtt cgaccgcgga tgacagacgg
1551 tgaggtttg gagaacatca tgccggaaga tgcaccatat catgacattt
1601 gcaagttcat tttgaggctc acatcagaaa gtgtaacttt ctcaggagag
1651 gggccagaag ttagtgattt gaacgaaaaa tggggaacaa ttgtgtacta
1701 tgagaaaaat ttgcaaattg gcgagaaaaa atgttcgaga ggaaatttcc
1751 acgtggatgg cggattcatt tgctctgaga atcgttacag tctcggactt
1801 gagccaaatc caattagaga accagtggcg ttaaagttc gtaaagcaat
1851 agtggatgga attcgcttt cctacaaaaa agacgggagt gtttggcttc
1901 aaaacgcat gaagtacccg gtattgtca cttctgggta tctcgacgag
1951 caatcaggag gcctaaagaa ggataaagtg cacaaagtt acggatgtgc
2001 gtatatcaaa acgttggct tcaacgttc caaacaaatc atcagagacg
2051 cgcttcttc caagcaaatg gcaacaatgt acttgcaagg aaaattgact
```

FIG. 11A
(Sheet 1)

2101 ccgatgaatt atatc. ga gaagaagact caggaagagc tgcgaaggga
2151 agcaacacgc accactgatt catggccaa gtactgttgt gtccgtgtct
2201 cgttctgcaa aggattggg gaagcatacc cagaacgccc gtcaattcat
2251 gattgtccag ttggattga gtgaaaatc aacattgcct acgatttcat
2301 ggattcaatc tgccagtaca taaccaactg cttcgagccg ctaggaatgg
2351 aagattttgc aaaattggga atcaacgtca gtgatgacta aatgataact
2401 ttttcactc accctactag atactgattt agtcttattc caaatcatcc
2451 aacgatatca aacttttcc ttgaacttt gcatactatg ttatcacaag
2501 ttccaagcag ttcaatatca aacataggat atgttaacaa ctttgataa
2551 gaatcaagtt accaactgtt cattgtgagc ttgagctgt atagaaggac
2601 aatgtatccc atacctcaat cttaatagt catcagtcac tggtcccgca
2651 ccaattttt cgattcgcat atgtcatata ttgcaccgtg gccctttta
2701 ttgtaacttt taatatattt tcttcccaac ttgtgaatat gattgatgaa
2751 ccaccattt gagtaataaa tgattttttt gtgg

FIG. 11A

```
   1 gtaatcaaat tgtaaaggaa aaaatttaat agtcagagta cacataaatg
  51 ggtgatcatc ataatttaac ggggccttcc ggtaccteca tcccgccaca
 101 gttcaactat tctcagcccg gtaccagcac cggaggcccg ctttatggtg
 151 gaaaaccttc tcatggattg gaagataatc ctgatgtaga ggaatatgag
 201 aggaaccctgc tcggggctgg agcaggtttt aatctgctca atgtaggaaa
 251 tatggctaat gttcccgacg agcacacacc gatgatgtca ccagtgaata
 301 caactacaaa gattctacaa cggagtggta ttaaaatgga aatcccgcca
 351 tattggatc cagacagtca ggatgatgac ccggaagatg gtgtcaacta
 401 cccggatcca gattaattg acacaaaaaa cacaaatatg accgagtacg
 451 attggatgt gttgaagctt ggaaaaccag cagtagatga agcacggaaa
 501 aagatcgaag ttcccgacgc tagtgcgccg ccaaacaaaa ttgtagaata
 551 ttgatgtat tatagaacgt taaaagaaag tgaactcata caactgaatg
 601 cgtatcggac aaaacgaaat cgattatcgt tgaacttggt caaaaacaat
 651 attgatcgag agttcgacca aaaagcttgc gagtcccctgg tgaaaaaatt
 701 gaaggataag aagaatgatc tccagaacct gattgatgtg gttctttcaa
 751 aaggtacaaa atataccggt tgcattacaa ttccaaggac acttgatggc
 801 cggttacagg tccacggaag aaaaggtttc cctcacgtag tctatggcaa
 851 actgtggagg ttaatgaaa tgacaaaaaa cgaaacgcgt catgtggacc
 901 actgcaagca cgcattgaa atgaaaagtg acatggtatg cgtgaatccc
 951 tatcactacg aaattgtcat tggaactatg attgttgggc agagggatca
1001 tgacaatcga gatatgccgc cgccacatca acgctaccac actccaggtc
1051 ggcaggatcc agttgacgat atgagtagat ttataccacc agcttccatt
1101 cgtccgcctc cgatgaacat gcacacaagg cctcagccta tgcctcaaca
1151 attgccttca gttggcgcaa cgttgcccca tcctctccca catcaggcgc
1201 cacataaccc aggggtttca catccgtact ccattgctcc acagacccat
1251 tacccgttga acatgaaccc aattccgcaa atgccgcaaa tgccacaaat
1301 gccaccacct ctccatcagg gatatggaat gaatgggccg agttgctctt
1351 cagaaaacaa caatccattc caccaaaatc accattataa tgatattagc
1401 catccaaatc actattccta cgactgtggt ccgaacttgt acgggtttcc
1451 aactccttat ccggattttc accatccttt caatcagcaa ccacaccagc
1501 cgccacaact atcacaaaac catacgtccc aacaaggcag tcatcaacca
1551 gggcaccaag gtcaggtacc gaatgatcca ccaattttcaa gaccagtgtt
1601 acaaccatca acagtcaacct tggacgtgtt ccgtcggtac tgtagacaga
1651 cattggaaa tcgattttt gaaggagaaa gtgaacaatc cggcgcaata
1701 attcggtcta gtaacaaatt cattgaagaa tttgattcgc cgatttgtgg
1751 tgtgacagtt gttcgaccgc ggatgacaga cggtgaggtt ttggagaaca
1801 tcatgccgga agatgcacca tatcatgaca tttgcaagtt cattttgagg
1851 ctcacatcag aaagtgtaac ttctcagga gaggggccag aagttagtga
1901 ttgaacgaa aaatgggga caattgtgta ctatgagaaa aattgcaaa
1951 ttggcgagaa aaaatgttcg agaggaaatt tccacgtgga tggcggattc
2001 attgctctg agaatcgtta cagtctcgga cttgagccaa atccaattag
2051 agaaccagtg gtgttaaag ttcgtaaagc aatagtggat ggaattcgct
```

FIG. 11B
(Sheet 1)

2101 tttcctacaa aaaag...ggg agtgtttggc ttcaaaaccg catgaagtac
2151 ccggtatttg tcacttctgg gtatctcgac gagcaatcag gaggcctaaa
2201 gaaggataaa gtgcacaaag tttacggatg tgcgtctatc aaaacgttcg
2251 gcttcaacgt ttccaaacaa atcatcagag acgcgcttct ttccaagcaa
2301 atggcaacaa tgtacttgca aggaaaattg actccgatga attatatcta
2351 cgagaagaag actcaggaag agctgcgaag ggaagcaaca cgcaccactg
2401 attcatggcc caagtactgt tgtgtccgtg tctcgttctg caaaggattt
2451 ggagaagcat acccagaacg cccgtcaatt catgattgtc cagttggat
2501 tgagttgaaa atcaacattg cctacgattt catggattca atctgccagt
2551 acataaccaa ctgcttcgag ccgctaggaa tggaagattt tgcaaaattg
2601 ggaatcaacg tcagtgatga ctaaatgata acttttttca ctcaccctac
2651 tagatactga ttagtctta ttccaaatca tccaacgata tcaaacttt
2701 tccttgaac ttgcatact atgttatcac aagttccaag cagttcaat
2751 acaaacatag gatatgtaa caacttttga taagaatcaa gttaccaact
2801 gttcattgtg agcttgagc tgtatagaag gacaatgtat ccctaccctc
2851 aatcttaat agtcatcagt cactggtccc gcaccaattt ttcgatteg
2901 catatgtcat atattgcacc gtggccctt ttattgtaac ttttaatata
2951 ttttcttccc aacttgtgaa tatgattgat gaaccaccat ttgagtaat
3001 aaatgtattt ttgtgg

FIG. 11B
(Sheet 2)

```
   1 gtaatcaaat tgtaaaggaa aaatattaat agtcagagta cacataaatg
  51 ggtgatcatc ataattaaac gggccttccc ggtaccctca tcccgccaca
 101 gttcaactat tctcagcccg gtaccagcac cggaggcccg ctttatggtg
 151 gaaaaccttc tcatggattg gaagataatc ctgatgtaga ggaatatgag
 201 aggaacctgc tcggggctgg agcaggtttt aatctgctca atgtaggaaa
 251 tatggctaat gaattaaac caataatcac attggacacg aaaccaccctc
 301 gtgatgccaa caagtcattg gcattcaatg gcgggttgaa gctaatcact
 351 ccgaaaactg aagttcccga cgagcacaca ccgatgatgt caccagtgaa
 401 tacaactaca aagattctac aacggagtgg tattaaaatg gaaatcccgc
 451 catattgga tccagacagt caggatgatg acccggaaga tggtgtcaac
 501 tacccggatc cagattatt tgacacaaaa aacacaaata tgaccgagta
 551 cgattggat gtgttgaagc ttggaaaacc agcagtagat gaagcacgga
 601 aaaagatcga agttcccgac gctagtgcgc cgccaaacaa aattgtagaa
 651 tattgatgt attatgaac gttaaaagaa agtgaactca tacaactgaa
 701 tgcgtatcgg acaaaacgaa atcgattatc gttgaacttg gtcaaaaaca
 751 atattgatcg agagttcgac caaaaagctt gcgagtccct ggtgaaaaaa
 801 ttgaaggata agaagaatga tctccagaac ctgattgatg tggttctttc
 851 aaaaggtaca aaatataccg gttgcattac aattccaagg acacttgatg
 901 gccggttaca ggtccacgga agaaaaggtt tccctcacgt agtctatggc
 951 aaactgtgga ggtttaatga aatgacaaaa aacgaaacgc gtcatgtgga
1001 ccactgcaag cacgcatttg aaatgaaaag tgacatggta tgcgtgaatc
1051 cctatcacta cgaaattgtc attggaacta tgattgttgg gcagagggat
1101 catgacaatc gagatatgcc gccgccacat caacgctacc acactccagg
1151 tcggcaggat ccagttgacg atatgagtag atttatacca ccagcttcca
1201 ttcgtccgcc tccgatgaac atgcacacaa ggcctcagcc tatgcctcaa
1251 caattgcctt cagttggcgc aacgtttgcc catcctctcc cacatcaggc
1301 gccacataac ccaggggttt cacatccgta ctccattgct ccacagaccc
1351 attacccgtt gaacatgaac ccaattccgc aaatgccgca aatgccacaa
1401 atgccaccac ctctccatca gggatatgga atgaatgggc cgagttgctc
1451 ttcagaaaac aacaatccat tccaccaaaa tcaccattat aatgatatta
1501 gccatccaaa tcactattcc tacgactgtg gtccgaactt gtacgggttt
1551 ccaactcctt atccggattt tcaccatcct ttcaatcagc aaccacacca
1601 gccgccacaa ctatcacaaa accatacgtc ccaacaaggc agtcatcaac
1651 cagggcacca aggtcaggta ccgaatgatc caccaatttc aagaccagtg
1701 ttacaaccat caacagtcac cttggacgtg ttccgtcggt actgtagaca
1751 gacattggga aatcgatttt ttgaaggaga aagtgaacaa tccggcgcaa
1801 taattcggtc tagtaacaaa ttcattgaag aatttgattc gccgatttgt
1851 ggtgtgacag ttgttcgacc gcggatgaca gacggtgagg ttttggagaa
1901 catcatgccg gaagatgcac catatcatga catttgcaag ttcattttga
1951 ggctcacatc agaaagtgta acttctcag gagaggggcc agaagttagt
2001 gattgaacg aaaaatgggg aacaattgtg tactatgaga aaaattgca
2051 aattggcgag aaaaaatgtt cgagaggaaa ttccacgtg gatggcggat
```

FIG. 11C
(Sheet 1)

2101 tcamgctc tgagaa tacagtctcg gacttgagcc aaatccaaat
2151 agaggaccag tggcgmaa agmcgtaaa gcaatagtgg atggaatcg
2201 cmmcctac aaaaagacg ggagtgmg gcttcaaaac cgcatgaagt
2251 accccggtat tgtcactct gggtatctcg acgagcaatc aggaggccta
2301 aggaaggata agtgccacaa agmtacgga tgtgcgtcta tcaaaacgtt
2351 tggcmcaac gmtccaaac aaatcatcag agacgcgcct cmtccaagc
2401 aaatggcaac aatgtacttg caaggaaaat tgactccgat gaatatatc
2451 tacgagaaga agactcagga agagctgcga agggaagcaa cacgcaccac
2501 tgatcattg gccaagtact gtgtgtccg tgtctcgttc tgcaaaggat
2551 tggagaagc ataccagaa cgcccgtcaa ttcatgattg tccagttgg
2601 attgagtga aaatcaacat tgcctacgat ttcatggatt caatctgcca
2651 gtacataacc aactgcttcg agccgctagg aatggaagat ttgcaaaat
2701 tgggaatcaa cgtcagtgat gactaaatga taacttttt cactcaccct
2751 actagatact gattagtct tatccaaat catccaacga tatcaaact
2801 ttccttga actgcata ctatgtatc acaagttcca agcagttca
2851 atacaaacat aggatatgt aacaacttt gataagaatc aagtaccaa
2901 ctgttcattg tgagcttga gctgtataga aggacaatgt atcccatacc
2951 tcaatctta atagtcatca gtcactggtc ccgcaccaat ttttcgatt
3001 cgcatatgtc atatattgca ccgtggccct ttttattgta actttaata
3051 tatttctc ccaacttgtg aatatgattg atgaaccacc atttgagta
3101 ataaatgtat ttttgtgg

FIG. 11C
(Sheet 2)

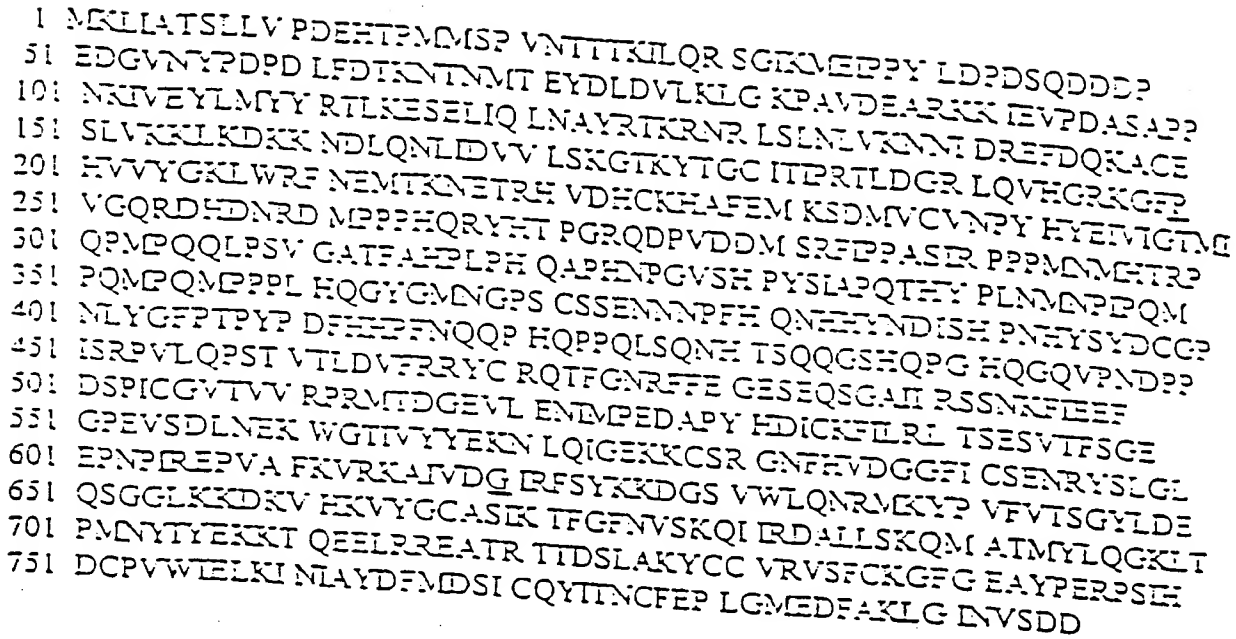

1 MKLLATSLLV PDEHTPMMSP VNTTTKILQR SGIKVEPPY LDPDSQDDDP
51 EDGVNYPDPD LFDTKNTNMT EYDLDVLKLG KPAVDEARKK IEVPDASAPP
101 NKIVEYLMYY RTLKESELIQ LNAYRTKRNR LSLNLVKNMI DREFDQKACE
151 SLVKKLKDKK NDLQNLDVV LSKGTKYTGC ITPRTLDGR LQVHGRKGFP
201 HVVYGKLWRF NEMTKNETRH VDHCKHAFEM KSDMVCVNPY HYEIVIGTMI
251 VGQRDHDNRD MPPPHQRYHT PGRQDPVDDM SRFPPASER PPPMNMHTRP
301 QPMPQQLPSV GATFAHPLPH QAPHNPGVSH PYSLAPQTHY PLNMNPPQM
351 PQMPQMPPPL HQGYGMNGPS CSSENNNPFH QNHHYNDISH PNHYSYDCGP
401 NLYGFPTPYP DFHHPFNQQP HQPPQLSQNH TSQQGSHQPG HQGQVPNDPP
451 ISRPVLQPST VTLDVFRRYC RQTFGNRFFE GESEQSGAII RSSNKFEEF
501 DSPICGVTVV RPRMTDGEVL ENEMPEDAPY HDICKFELRL TSESVTFSGE
551 GPEVSDLNEK WGTIVYYEKN LQIGEKKCSR GNFHVDGGFI CSENRYSLGL
601 EPNPEREPVA FKVRKAIVDG RFSYKKDGS VWLQNRMKYP VFVTSGYLDE
651 QSGGLKKDKV HKVYGCASEK TFGFNVSKQI RDALLSKQM ATMYLQGKLT
701 PMNYTYEKKT QEELRREATR TTDSLAKYCC VRVSFCKGFG EAYPERPSEH
751 DCPVWIELKI NLAYDFMDSI CQYTINCFEP LGMEDFAKLG INVSDD

FIG. 12A

1 MGDEH NLTGLPGTSI PPQFNYSQPG TSTGGPLYGG
51 KPSHGLEDP DVEEYERNLL GAGAGFNLLN VGNMANVPDE HTPMMSPVNT
101 TTKILQRSGI KMEPPYLDP DSQDDDPEDG VNYPDPDLFD TKNTNMTEYD
151 LDVLKLGKPA VDEARKKIEV PDASAPPNKI VEYLMYYRTL KESELIQLNA
201 YRTKRNRLSL NLVKNNDRE FDQKACESLV KKLKDKKNDL QNLDVVLSK
251 GTKYTGCITI PRTLDGRLQV HGRKGFPHVV YGKLWRFNEM TKNETRHVDH
301 CKHAFEMKSD MVCVNPYHYE IVIGTMIVGQ RDHDNRDMPP PHQRYHTPGR
351 QDPVDDMSRF PPASRPPP MNNMHTRPQPM PQQLPSVGAT FAHPLPHQAP
401 HNPGVSHPYS LAPQTHYPLN MNPPQMPQM PQMPPPLHQG YGMNGPSCSS
451 ENNNPFHQNH HYNDISHPNH YSYDCGPNLY GFPTPYPDFH HPFNQQPHQP
501 PQLSQNHTSQ QGSHQPGHQG QVPNDPPISR PVLQPSTVTL DVFRRYCRQT
551 PGNRFFEGES EQSGAIRSS NKFEEFDSP ICGVTVVRPR MTDGEVLENI
601 MPEDAPYHDI CKFILRLTSE SVTFSGEGPE VSDLNEKWGT IVYYEKNLQI
651 GEKKCSRGNF HVDGGFICSE NRYSLGLEPN PREPVAFKV RKAIVDGRF
701 SYKKDGSVWL QNRMKYPVFV TSGYLDEQSG GLKKDKVFKV YGCASEKTFG
751 FNVSKQIRD ALLSKQMATM YLQGKLTPMN YTYEKKTQEE LRREATRTTD
801 SLAKYCCVRV SFCKGFGEAY PERPSEHDCP VWELKNLA YDFMDSICQY
851 ITNCFEPLGM EDFAKLGENV SDD

FIG. 12B

```
  1 MGDEH NLTGLPGTSI PPQFNYSQPG TSTGGPLYGG
 51 KPSHGLEDP DVEEYFRNLL GAGAGFNLLN VGNMANEFKP ETLDTKPPR
101 DANKSLAFNG GLKLITPKTE VPDEHTPMMS PVNTITTKEQ RSGLKMEPP
151 YLDPDSQDDD PEDGVNYPDP DLFDTKNTNM TEYDLDVLKL GKPAVDEARK
201 KIEVPDASAP PNKIVEYLMY YRTLKESELI QLNAYRTKRN RLSLNLVKNN
251 IDREFDQKAC ESLVKKLKDK KNDLQNLIDV VLSKGTKYTG CITPRTLDG
301 RLQVHGRKGF PHVVYGKLWR FNEMTKNETR HVDHCKHAFE MKSDMVCVNP
351 YHYEIVIGTM IVGQRDHDNR DMPPPHQRYH TPGRQDPVDD MSRFPPASI
401 RPPPMNMGTR PQPMPQQLPS VGATFAHPLP HQAPHNPGVS HPYSLAPQTH
451 YPLNMNPPQ MPQMPQMPPP LHQGYGMNGP SCSSENNNPF HQNFHYNDIS
501 HPNHYSYDCG PNLYGFPTPY PDFFHPFNQQ PHQPPQLSQN HTSQQGSHQP
551 GHQGQVPNDP PISRPVLQPS TVTLDVFRRY CRQTFGNRFF EGESEQSGAI
601 ERSSNKFEE FDSPICGVTV VRPRMTDGEV LENEMPEDAP YHDICKFILR
651 LTSESVTFSG EGPEVSDLNE KWGTIVYYEK NLQIGEKKCS RGNFHVDGGF
701 ICSENRYSLG LEPNPEREPV AFKVRKAIVD GERFSYKKDG SVWLQNRMKY
751 PVFVTSGYLD EQSGGLKKDK VHKVYGCASI KTFGFNVSKQ ERDALLSKQ
801 MATMYLQGKL TPMNYTYEKK TQEELRREAT RTTDSLAKYC CVRVSFCKGF
851 GEAYPERPSI HDCPVWIELK DNLAYDFMDS ICQYTINCFE PLGMEDFAKL
901 GENVSDD
```

FIG. 12C

Figure13A spliced form exons 123467891011

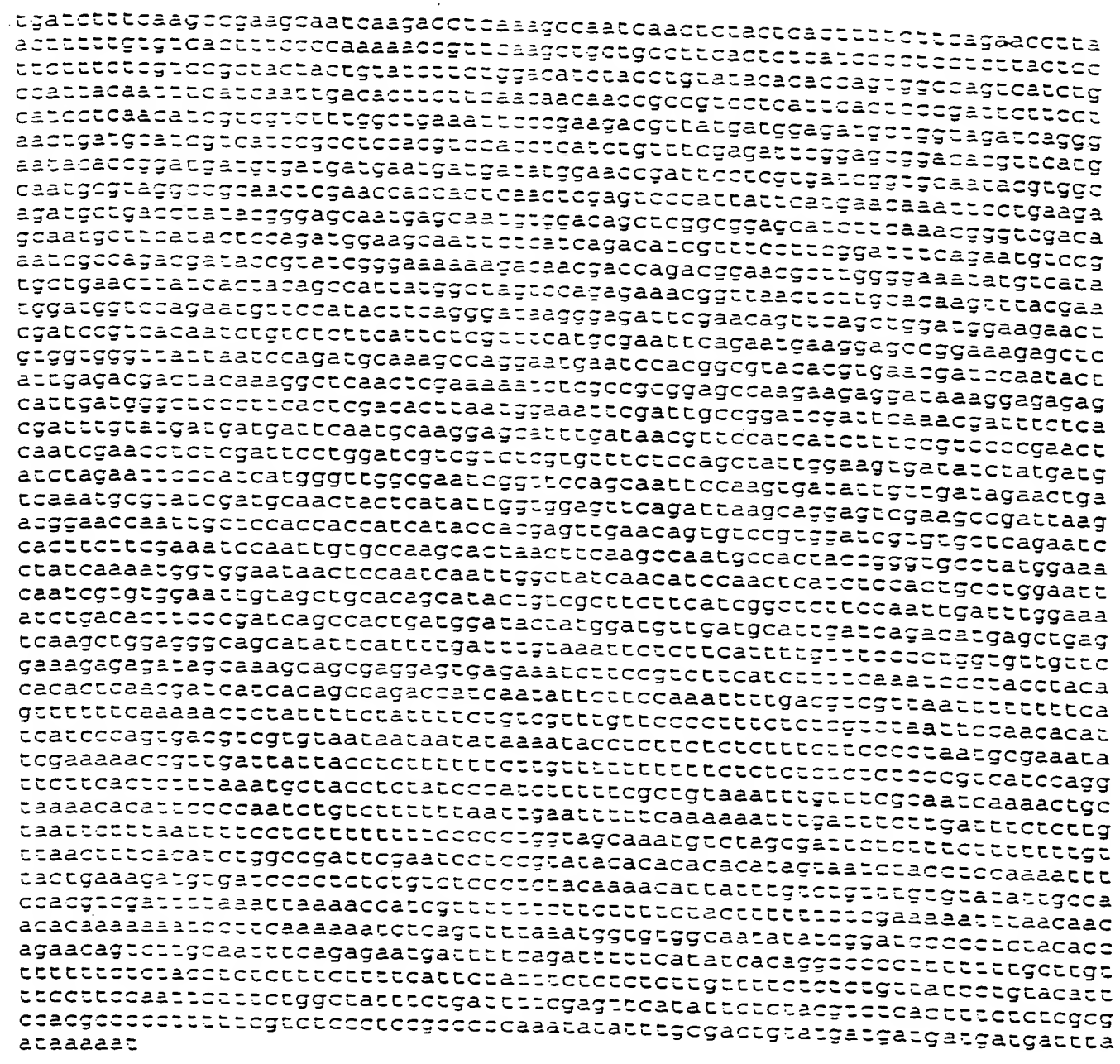


Figure 13B daf-16 spliced form 567891011

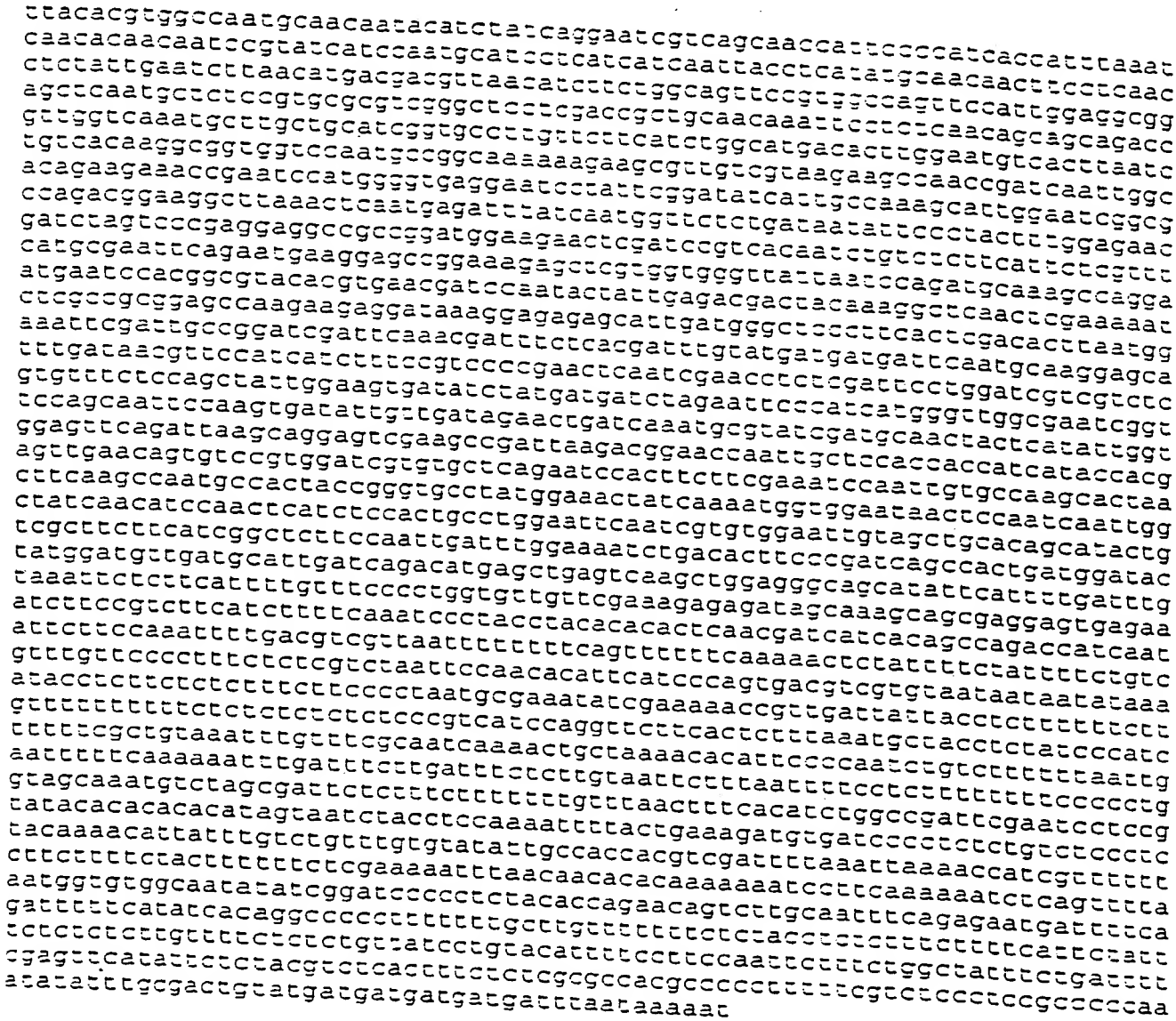


Figure 14A
DAF-16 exons 1234 67891011

MMEMLVDQGTDASSSASTSTSSVSRFGADTFMNTPDDVMMNDDMEPIPRDR
CNTWPMRRPQLEPPLNSSPIIHEQIPEEDADLYGSNEQCGQLGGASSNGST
AMLHTPDGSNSHQTSPPSDFRMSESPDDTVSGKKTTTPRNAWGNMSYAELI
TTAIMASPEKRLTLAQVYEMMVQNVPYFRDKGDSNSSAGWKNSIRHNLSLH
SRFMRIQNEGAGKSSWWVINPDAKPGMNPRRTRERSNTIETTTKAQLEKSR
RGAKKRIKERALMGSLHSTLNGNSIAGSIQTISHDLYDDDSMQGAFDNVPS
SFRPRTQSNLSIPGSSSRVSPAIGSDIYDDLEFPSWVGESVPAIPSDIVDR
TDQMRIDATTHIGGVQIKQESKPIKTEPIAPPPSYHELNSVRGSCAQNPLL
RNPIVPSTNFKPMPLPGAYGNYQNGGITPINWLSTSNSSPLPGIQSCGIVA
AQHTVASSSALPIDLENLTLPDQPLMDTMDVDALIRHELSQAGGQHIHFDL

Figure 14B
DAF-16 exons 567891011
MQQYIYQESSATIPHHHLNQHDNPYHPMHPHHQLPHMQQLPQPLLNLNMTT
LTSSGSSVASSIGGGAQCSPCASGSSTAATNSSQQQQTVGQMLAASVPCSS
SGMTLGMSLNLSQGGGPMPAKKKRCRKKPTDQLAQKKPNPWGEESYSDIIA
KALESAPDGRLKLNEIYQWFSDNIPYFGERSSPEEAAGWKNSIRHNLSLHS
RFMRIQNEGAGKSSWWVINPDAKPGMNPRRTRERSNTIETTTKAQLEKSRR
GAKKRIKERALMGSLHSTLNGNSIAGSIQTISHDLYDDDSMQGAFDNVPSS
FRPRTQSNLSIPGSSSRVSPAIGSDIYDDLEFPSWVGESVPAIPSDIVDRT
DQMRIDATTHIGGVQIKQESKPIKTEPIAPPPSYHELNSVRGSCAQNPLLR
NPIVPSTNFKPMPLPGAYGNYQNGGITPINWLSTSNSSPLPGIQSCGIVAA
QHTVASSSALPIDLENLTLPDQPLMDTMDVDALIRHELSQAGGQHIHFDL

FIG. 15 (Sheet 1)

corrected age-1 cDNA and AGE-1 protein sequence

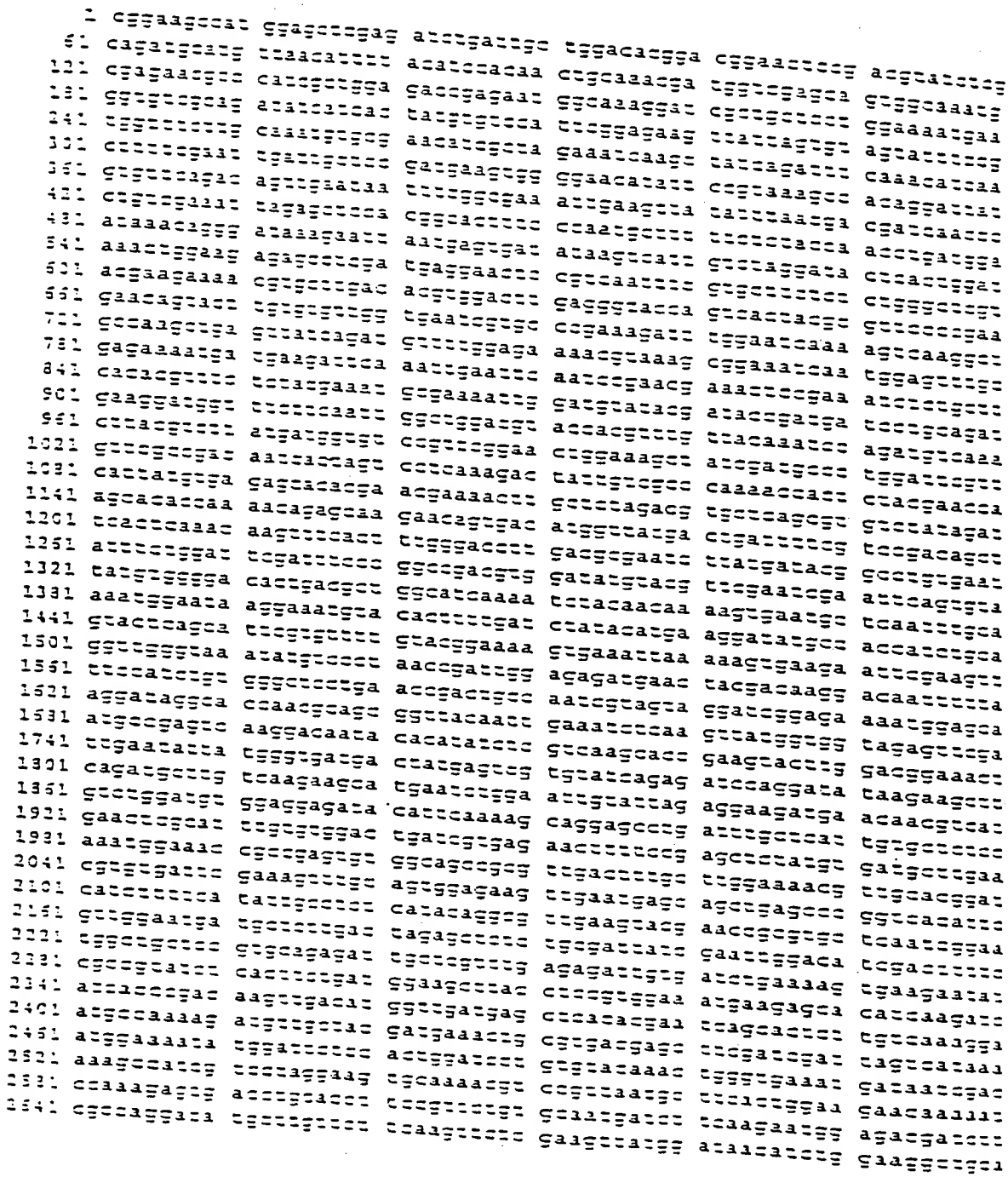


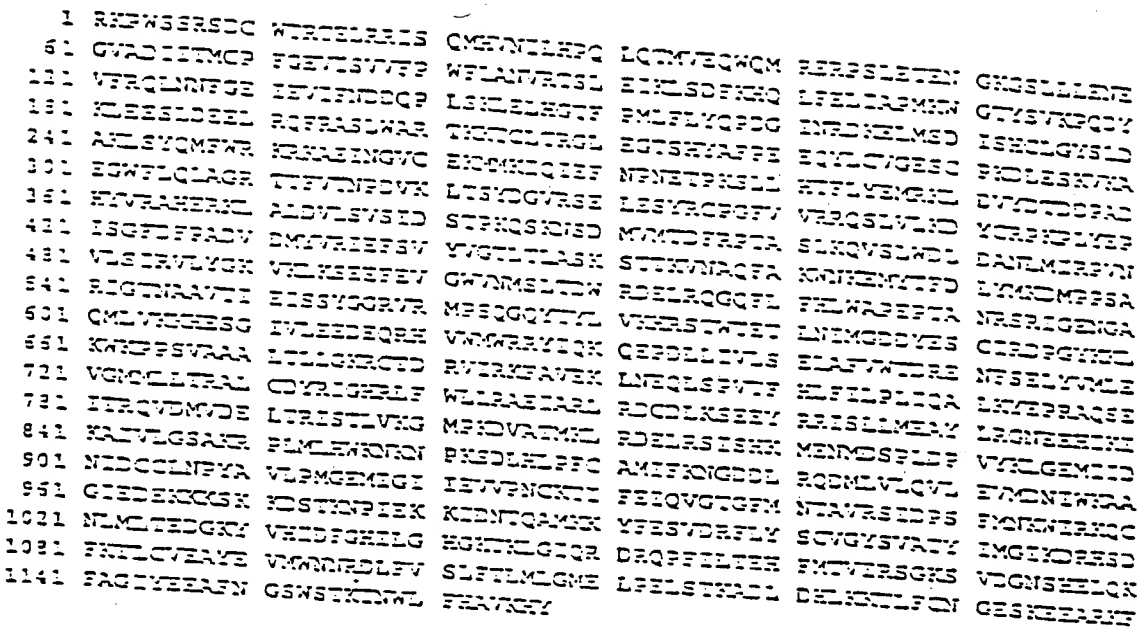
Fig. 15

```
2701 aatattgatt  :tgtttgaa cccgtacgca gttcttcaaa tgggagaaat gattggaatt
2761 attgaagttg tgcctaattg taaaacaata ttcgagattc aagttggaac aggattcatg
2821 aatacagcag ttcggagtat tgatccttcg ttctatgaata agtggatttg gaaacgatgc
2881 ggaattgaag atgaaaagaa gaaaagcaaa aaggacttta cgaaaaatcc catcgaaaag
2941 aagattgata atacttaagc catgaagaaa tattttgaaa gtgtcgatcg attcctatac
3001 tcgtgtgttg gatattcagt tgccacgtac ataatgggaa tcaaggatcg ttcacagtgat
3061 aattcgatgc tcacttgaaga tggaaaatat gtccacattg attcggttca catttttggga
3121 catggaaaga ctaaacttgg gatccagcga gatcgttcaac cgttttattct aaccgaatac
3181 tttatgacag tgatttgatc gggtaaattt gtggatggaa attcgcatga gttacaaaaa
3241 ttcaaaacgt tatgcgtcga agcctacgaa gtaatgtgga ataatcgaga ttttgttcgtt
3301 tccctgttca cattgatgct cggaatggag ttgccttgagc tgtcgacgaa agcggatttg
3361 gatcatttga agaaaaccct ctttgcaat ggagaaagca aagaagaagc gagaaagttt
3421 ttcgcttgaa tctacgaaga agccttcaat ggatcatggt ctaccaaaac gaattggctc
3481 ttccacgcag tcaaacatta ctga
```

29/53

FIG. 16

AGE-1 protein



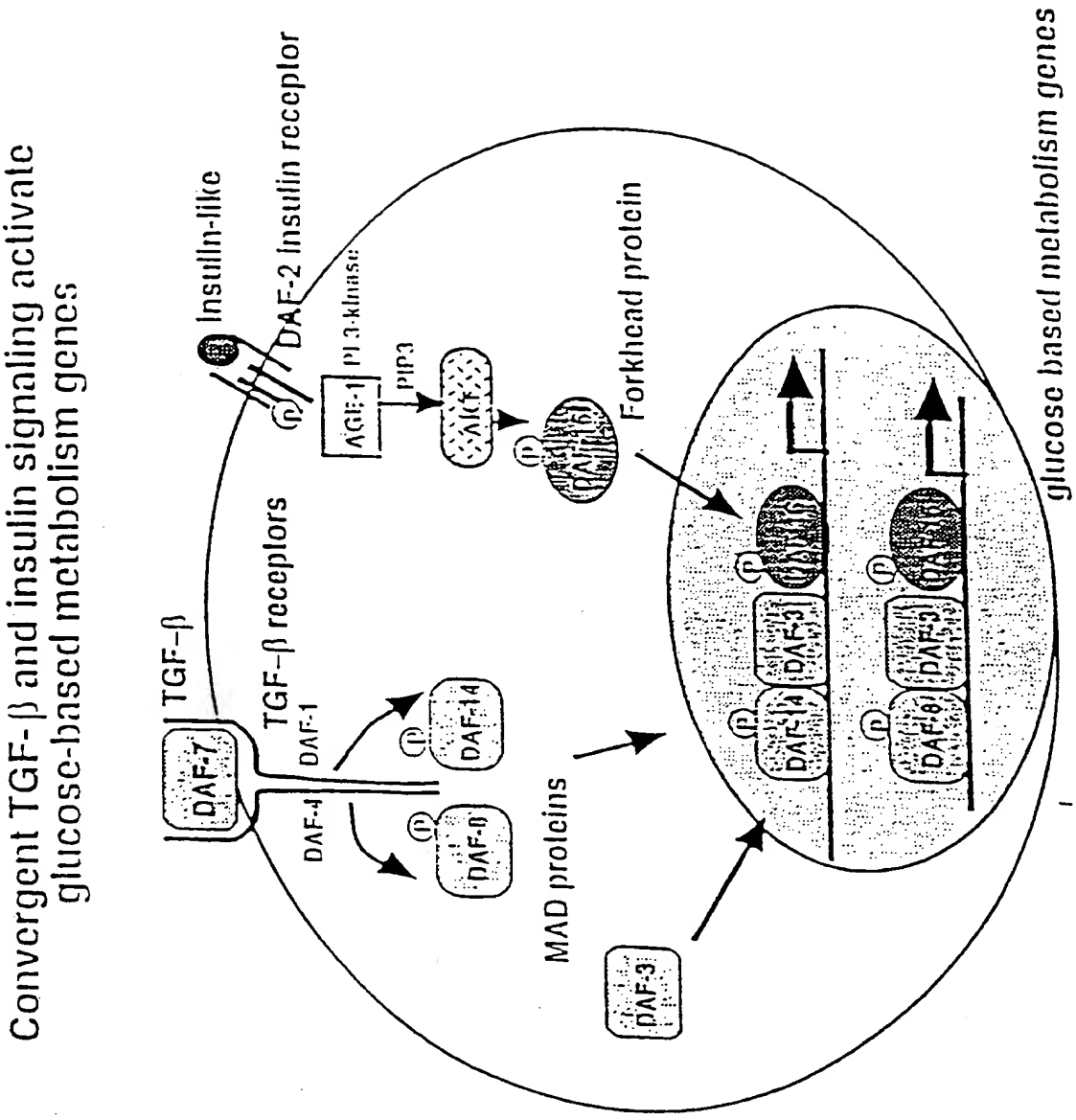


FIG 17

In pheromone, no TGFβ or insulin-like signals causes repression of anabolic genes

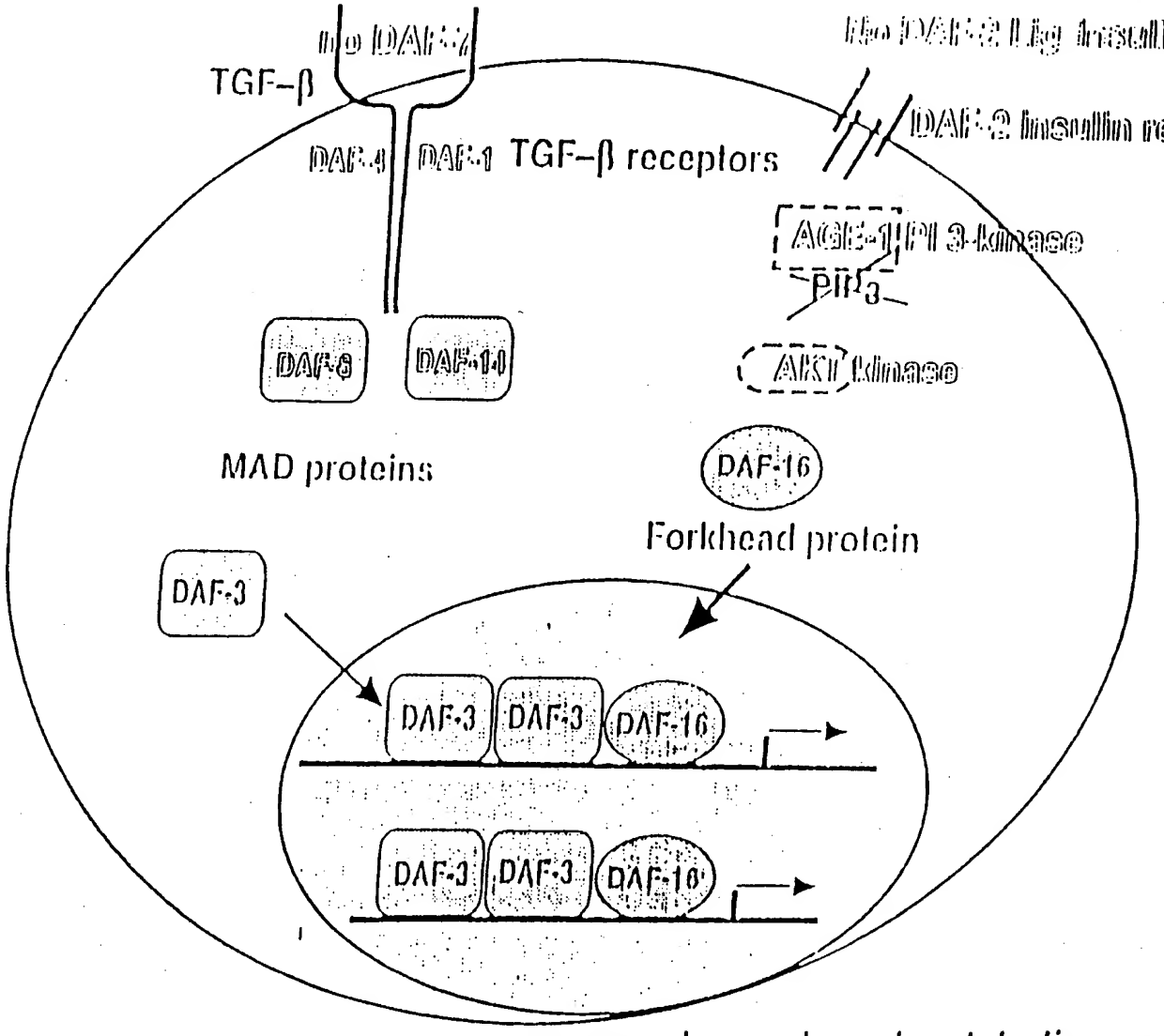
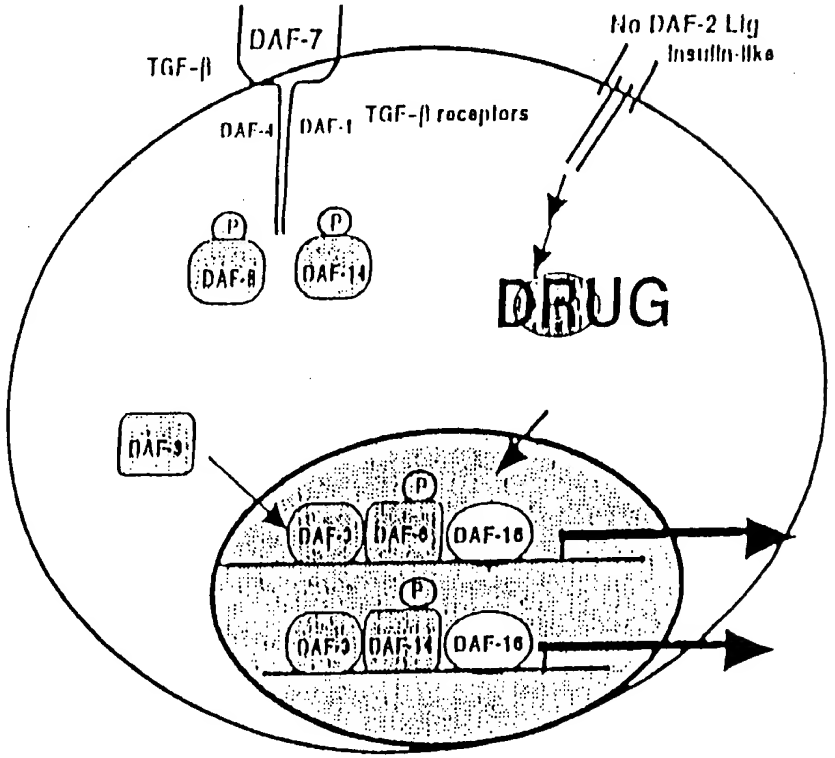


FIG. 18

repress glucose based metabolism genes

Drugs that inhibit DAF-16 or DAF-3
(or proteins in the pathway)
can be discovered using reporter genes bearing their cognate
binding sites



drug causes a decrease in DAF-16 activity, activating the reporter
gene like a daf-16 mutant.
This bypasses the need for insulin

FIG. 19

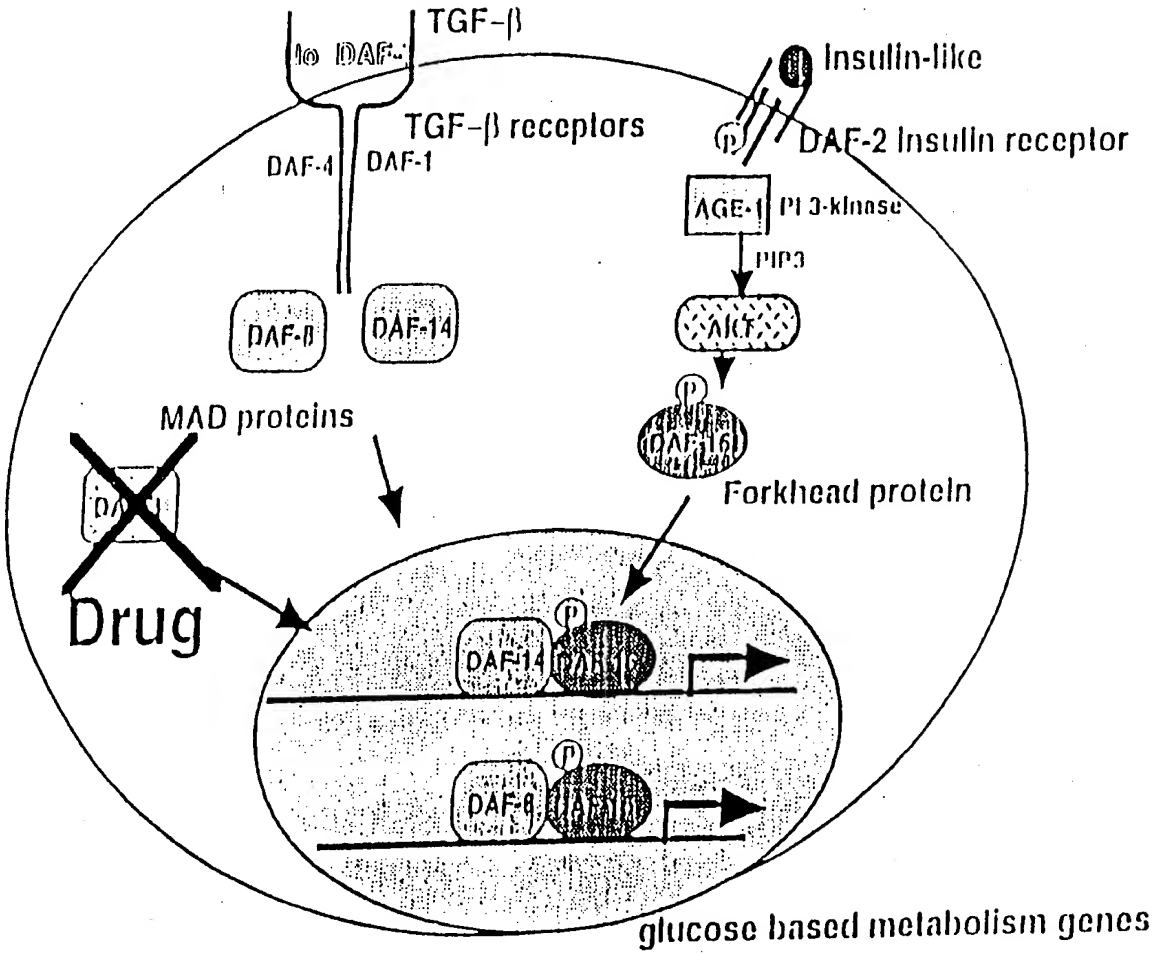


FIG. 20

34/53

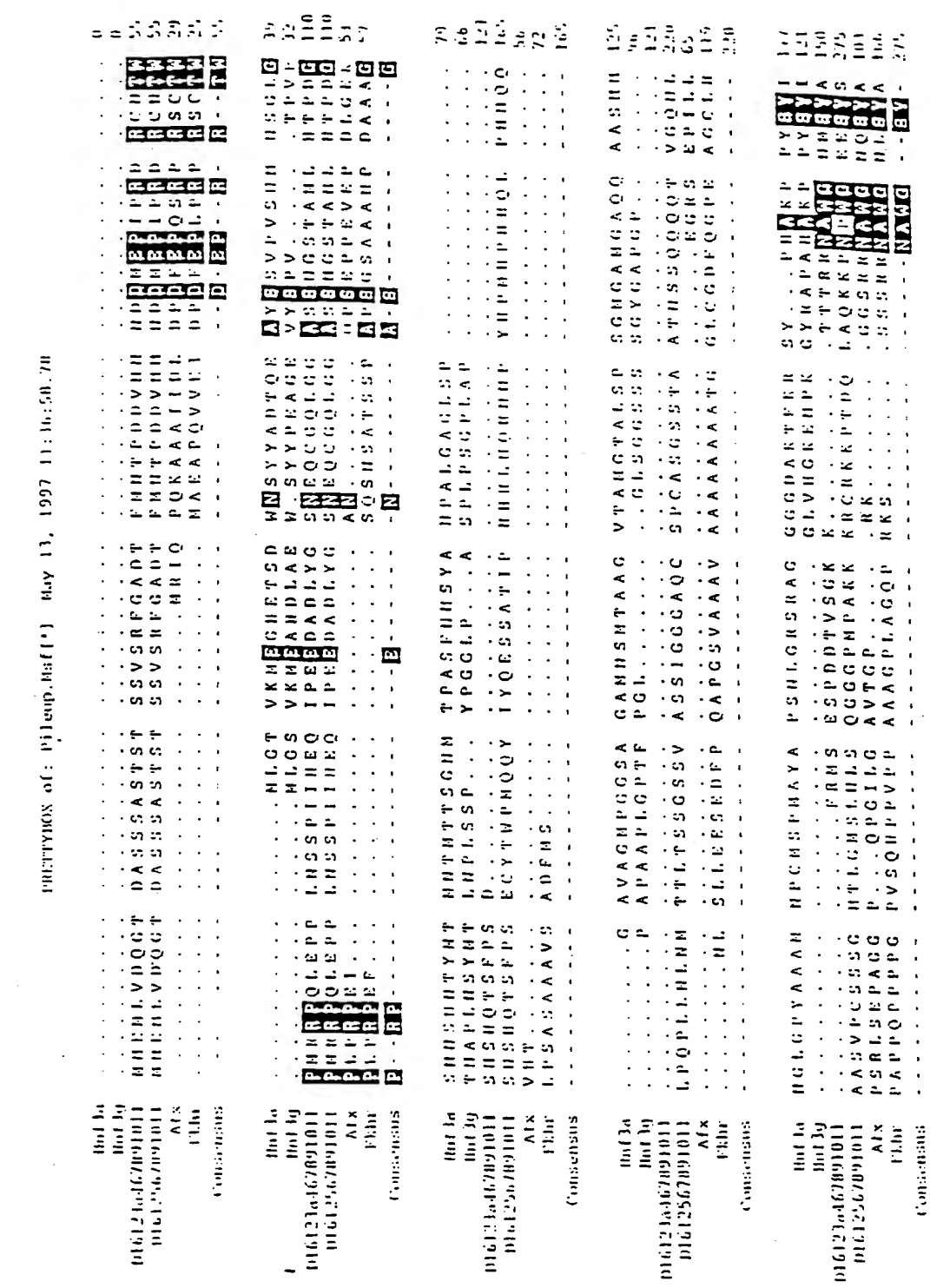


Fig. 2 A
Sheet 1 of 3

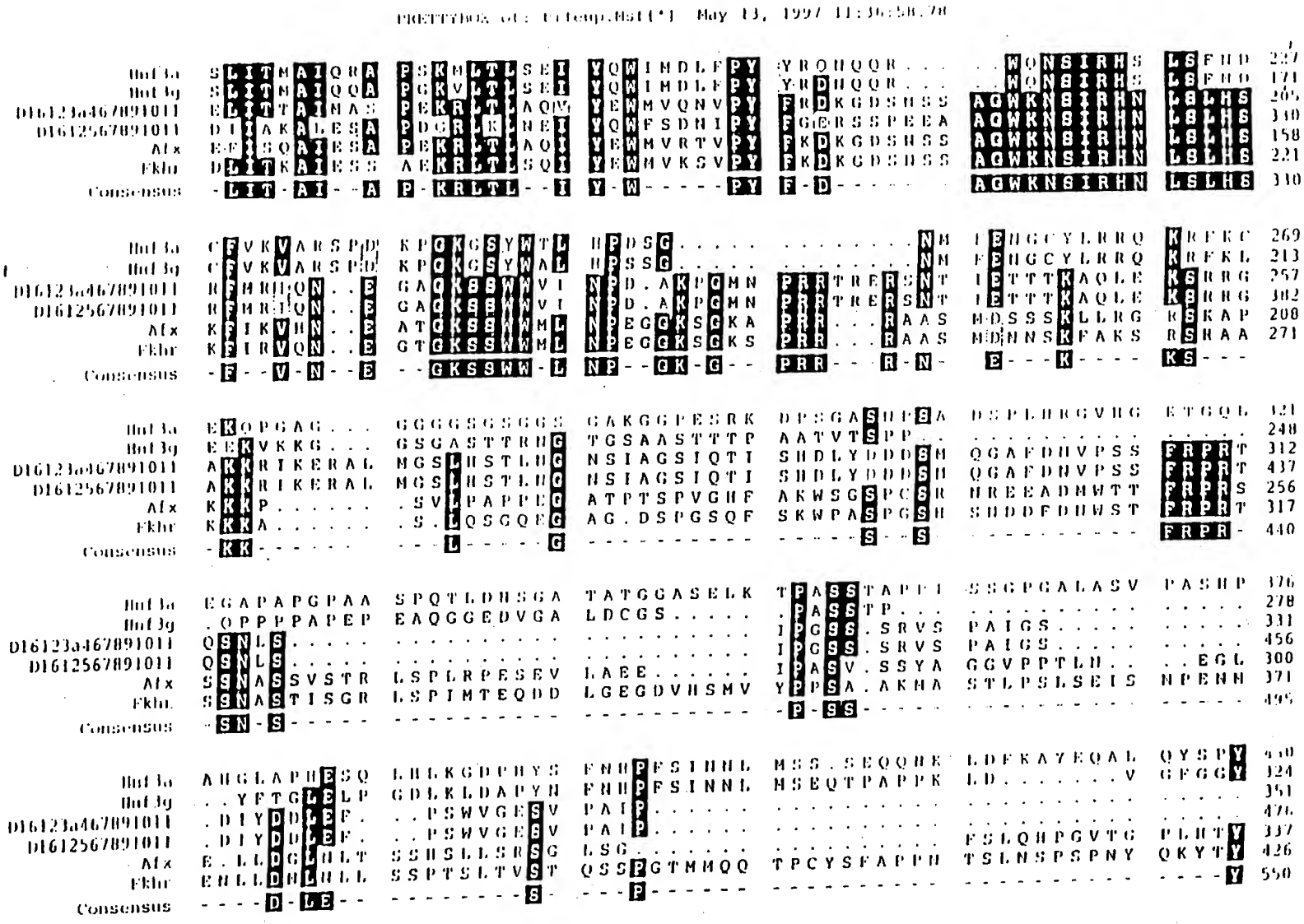


Fig. 21A

sheet 2 of 3

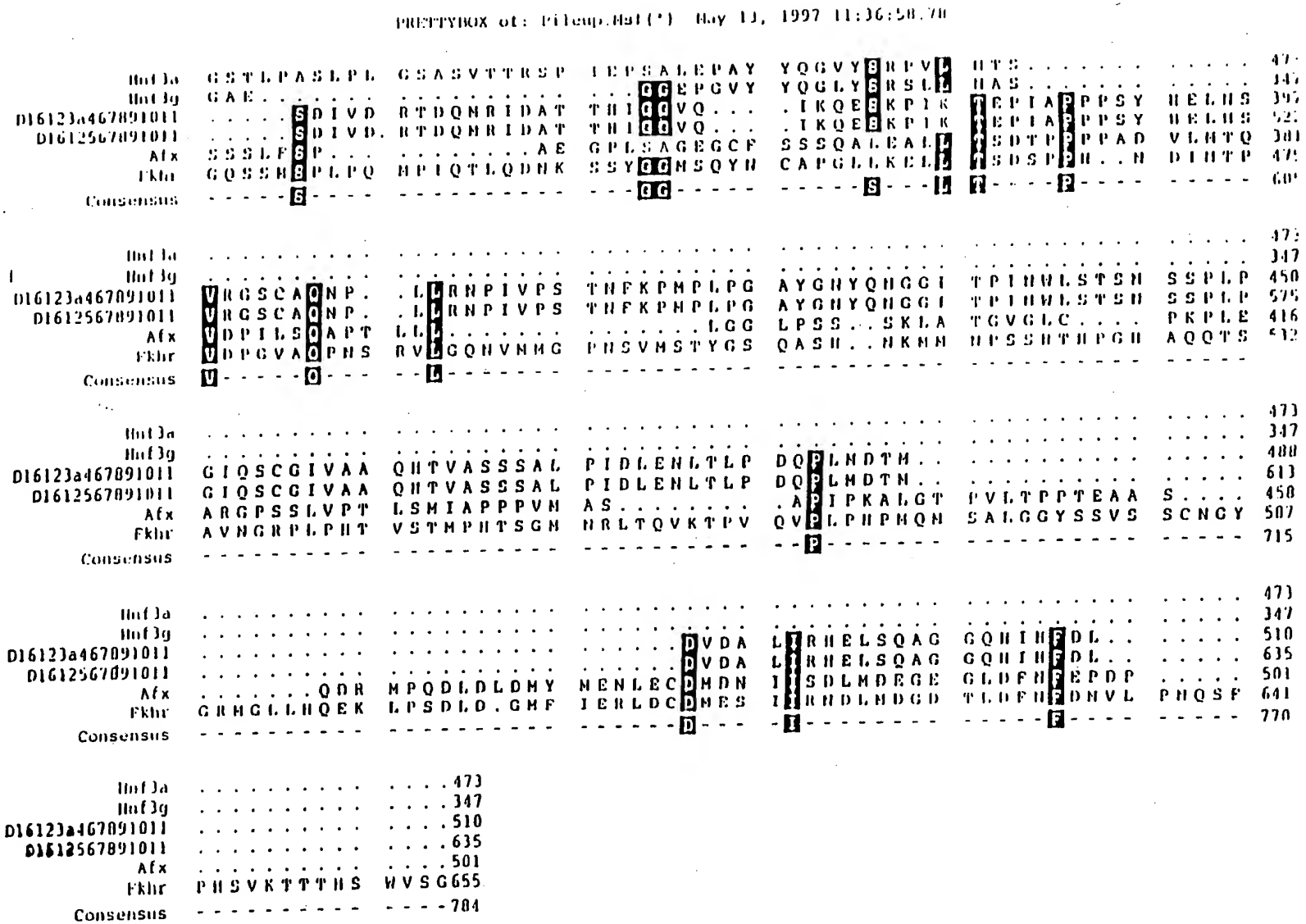


Fig 21A

Sheet 3 of 3

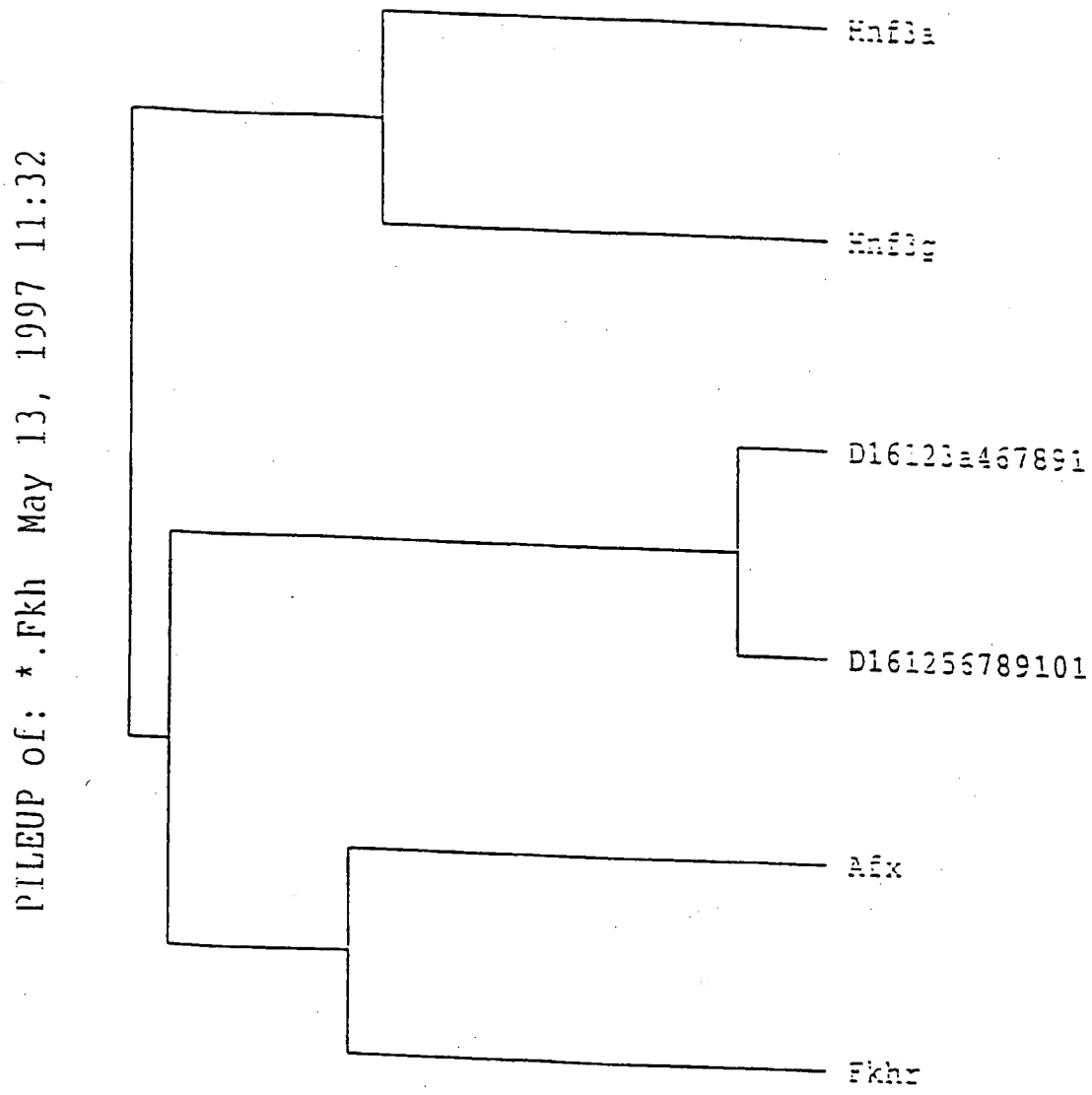

PILEUP of: *.Fkh May 13, 1997 11:32
Hnf3a
Hnf3g
D16123a467891
D161256789101
Afx
Fkhr

Fig. 21B

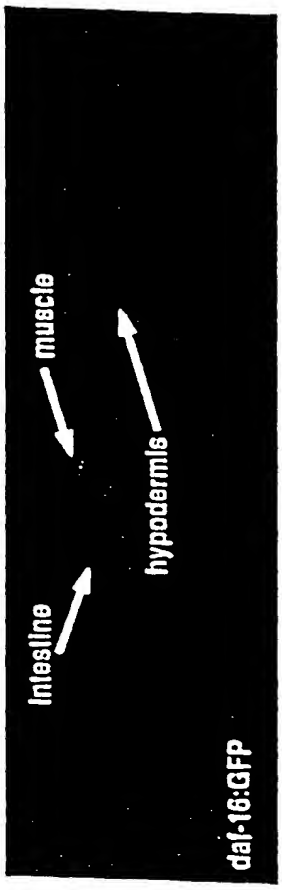
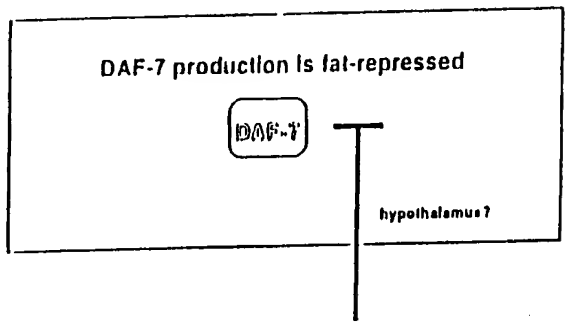


Fig. 22

Injection of DAF-7 bypasses obesity-induced defects in insulin-regulation of metabolism



fatty acids in blood repress DAF-7 in analogy to pheromone regulation of DAF-7 in C. elegans

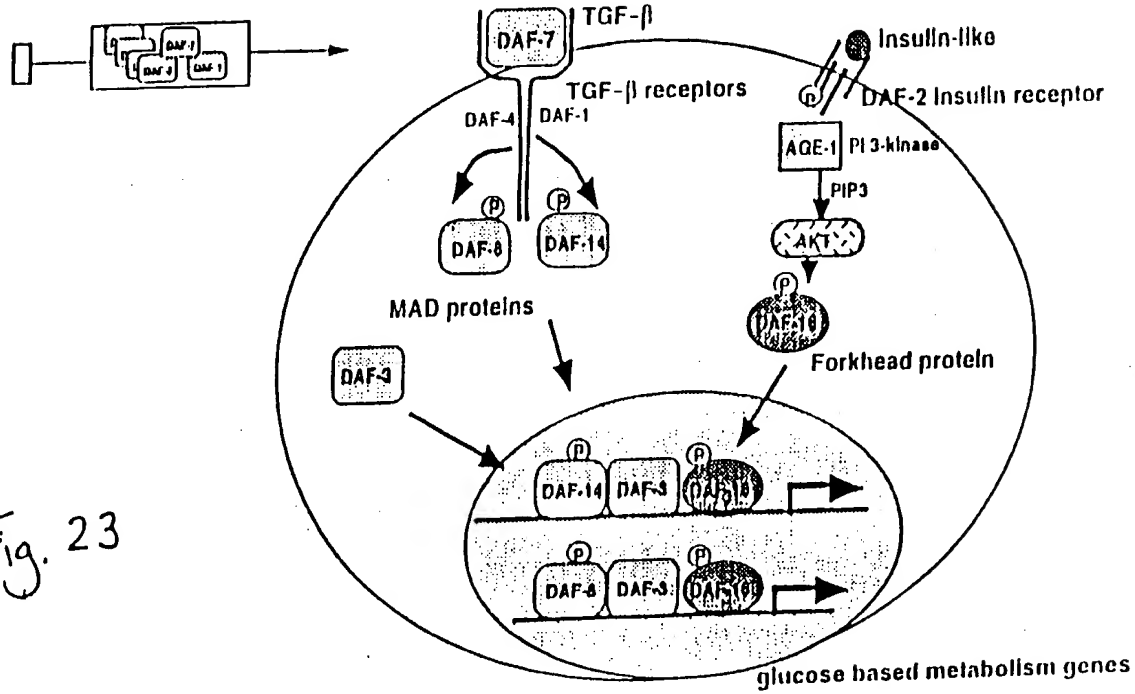


Fig. 23

Figure 25. comparison of the human AKT protein sequence to the cosmid sequence C12D8, located in the genetic interval where sup(mg144) maps. Numbering in the AKT protein sequence by amino acid residues, and in the cosmid sequence by nucleotide position.

```
 Score = 450 (207.4 bits), Expect = 5.2e-165, Sum P(7) = 5.2e-165
 Identities = 79/121 (65%), Positives = 97/121 (80%), Frame = -1

Query:    319 EVLEDNDYGRAVDWWGLGVVMYEMMCGRLPFYNQDHEKLFELILMEEIRFPRTLGPEAKS 378
              +VL+D-DYGR VDWWG-GVVMYEMMCGRLPFY++DH KLFELI+  ++RFP  L  EA++
Sbjct:  33685 QVLDDHDYGRCVDWWGVGVVMYEMMCGRLPFYSKDHNKLFELIMAGDLRFPSKLSQEART 33864

Query:    379 LLSGLLKKDPTQRLGGGSEDAKEIMQHRFFANIVWQDVYEKKLSPPFKPQVTSETDTRYF 438
              LL+GLL KDPTQRLGGG EDA EI +  FF  + W+  Y K-- PP-KP V SETDT YF
Sbjct:  33865 LLTGLLVKDPTQRLGGGPEDALEICRADFFRTVDWEATYRKEIEPPYKPNVQSETDTSYF 34044

Query:    439 D 439
              D
Sbjct:  34045 D 34047

 Score = 256 (118.0 bits), Expect = 5.2e-165, Sum P(7) = 5.2e-165
 Identities = 43/66 (72%), Positives = 59/66 (89%), Frame = +1

Query:    146 TMNEFEYLKLLGKGTFGKVILVKEKATGRYYAMKILKKEVIVAKDEVAHTLTENRVLQNS 205
              TM +F++LK-LGKGTFGKVIL KEK T + YA+KILKK+VI+A--EVAHTLTENRVLQ
Sbjct:  32314 TMEDFDFLKVLGKGTFGKVILCKEKRTQKLYAIKILKKDVIIAREEVAHTLTENRVLQRC 32493

Query:    206 RHPFLT 211
              +HPFLT
Sbjct:  32494 KHPFLT 32511

 Score = 190 (87.6 bits), Expect = 5.2e-165, Sum P(7) = 5.2e-165
 Identities = 36/45 (80%), Positives = 37/45 (82%), Frame = +2

Query:    276 KLENLMLDKDGHIKITDFGLCKEGIKDGATMKTFCGTPEYLAPEV 320
              KLENL+LDKDGHIKI DFGLCKE I  G    TFCGTPEYLAPEV
Sbjct:  33509 KLENLLLDKDGHIKIADFGLCKEEISFGDKTSTFCGTPEYLAPEV 33643

 Score = 138 (86.7 bits), Expect = 5.2e-165, Sum P(7) = 5.2e-165
 Identities = 37/57 (64%), Positives = 42/57 (73%), Frame = +3

Query:    209 FLTALKYSFQTHDRLCFVMEYANGGELFFHLSRERVFSEDRARFYGAEIVSALDYLH 265
              +   LKYSFQ    LCFVM++ANGGELF H+ +   FSE RARFYGAEIV AL YLH
Sbjct:  32667 YFQELKYSFQEQHYLCFVMQFANGGELFTHVRKCGTFSEPRARFYGAEIVLALGYLH 32837

 Score = 166 (75.5 bits), Expect = 5.2e-165, Sum P(7) = 5.2e-165
 Identities = 29/59 (49%), Positives = 42/59 (71%), Frame = +1

Query:     53 MNFSVAQCQLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREEWATAIQTVADGLK 111
              + F++   Q M  E-PRPN F++RCLQWTTVIERTF+ E+ E R+ W  AI++++   K
Sbjct:  31846 STFAIFYFQTMLFEKPRPNMFMVRCLQWTTVIERTFYAESAEVRQRWIHAIESISHKYK 32022

 Score = 114 (61.8 bits), Expect = 5.2e-167, Sum P(8) = 5.2e-167
 Identities = 24/33 (72%), Positives = 30/33 (90%), Frame = +3

Query:    210 LTALKYSFQTHDRLCFVMEYANGGELFFHLSRE 242
              L  LKYSFQT-DRLCFVME+A GG+L++HL+RE
Sbjct:  33156 LQELKYSFQTNDRLCFVMEFAIGGDLYYHLNRE 33254
```

Fig. 25

Mapping of *sug(. g144)* to the same genetic region as Akt

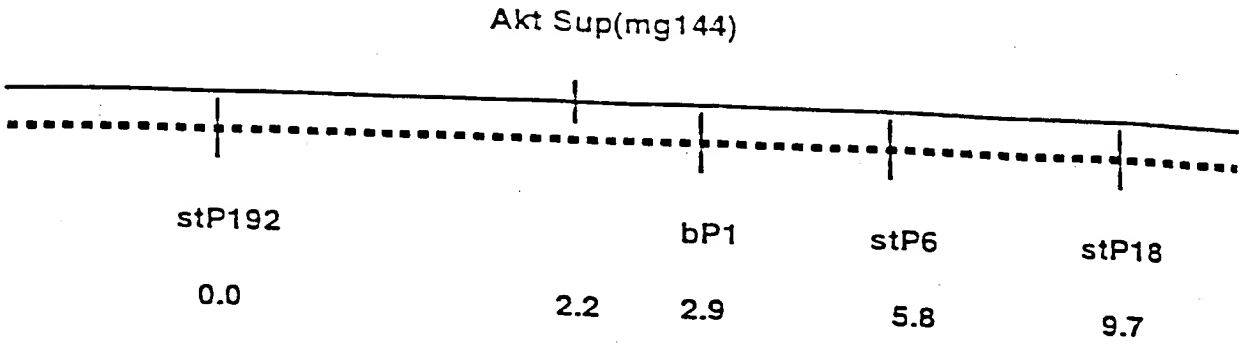


Fig. 24

Fig. 26A

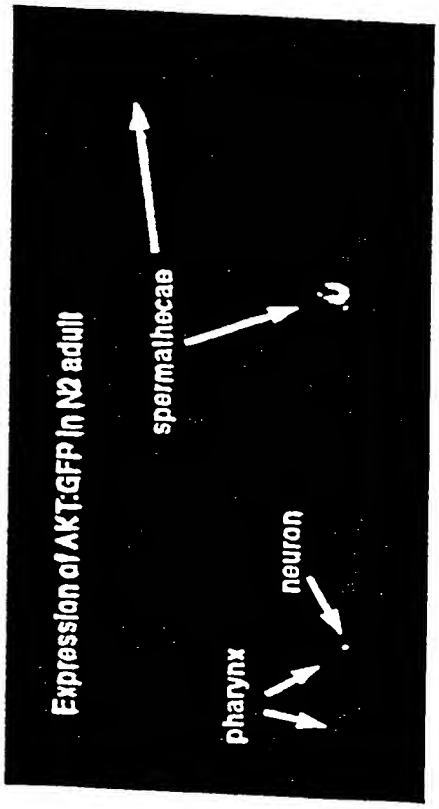

Expression of AKT:GFP in N2 adult
spermatheca
neuron
pharynx

Fig. 26B

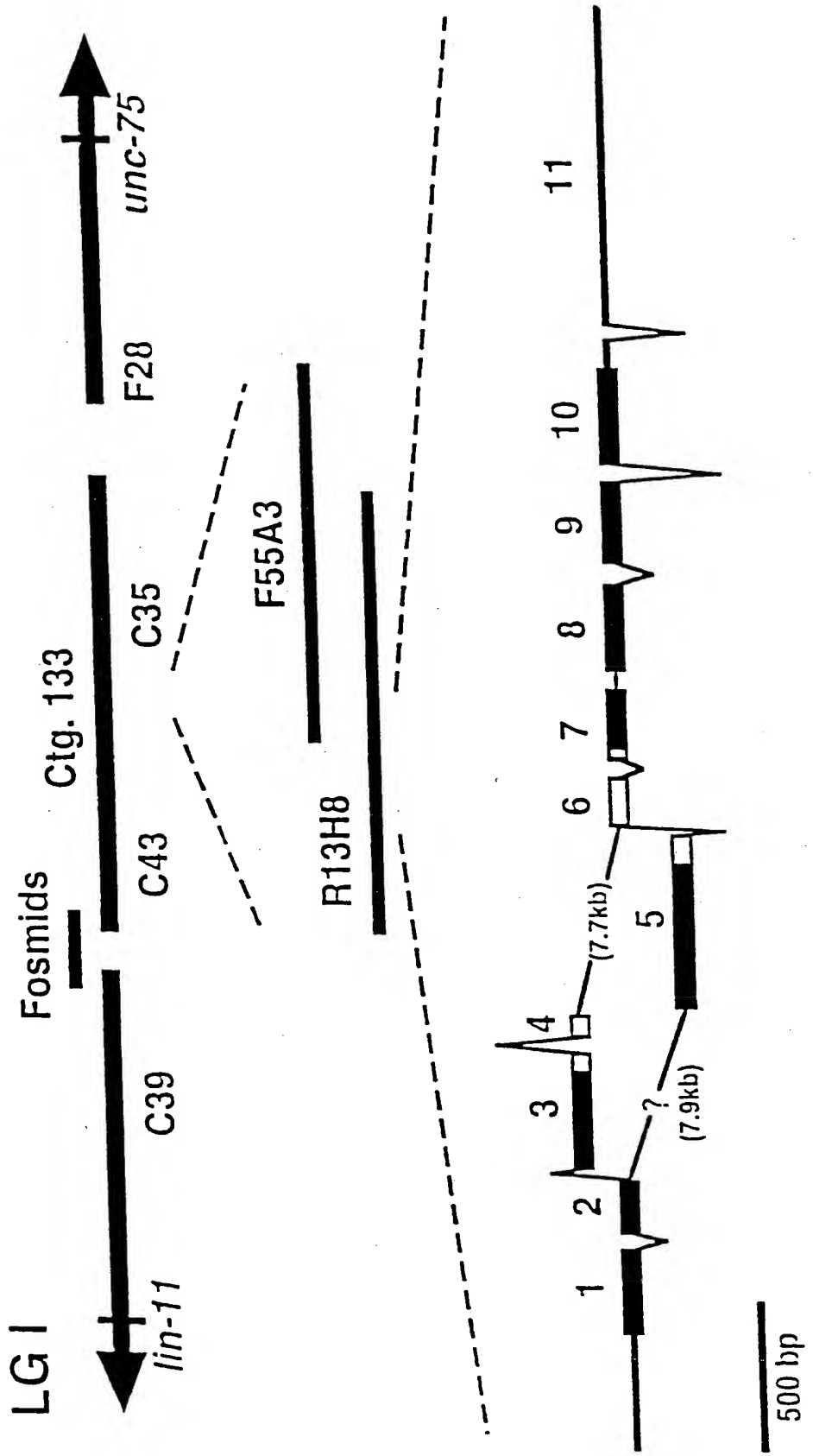

LG I
lin-11
unc-75
Fosmids
Ctg. 133
C39
C43
C35
F28
F55A3
R13H8
1
2
3
4
5
6
7
8
9
10
11
(7.7kb)
?
(7.9kb)
500 bp

Fig. 27

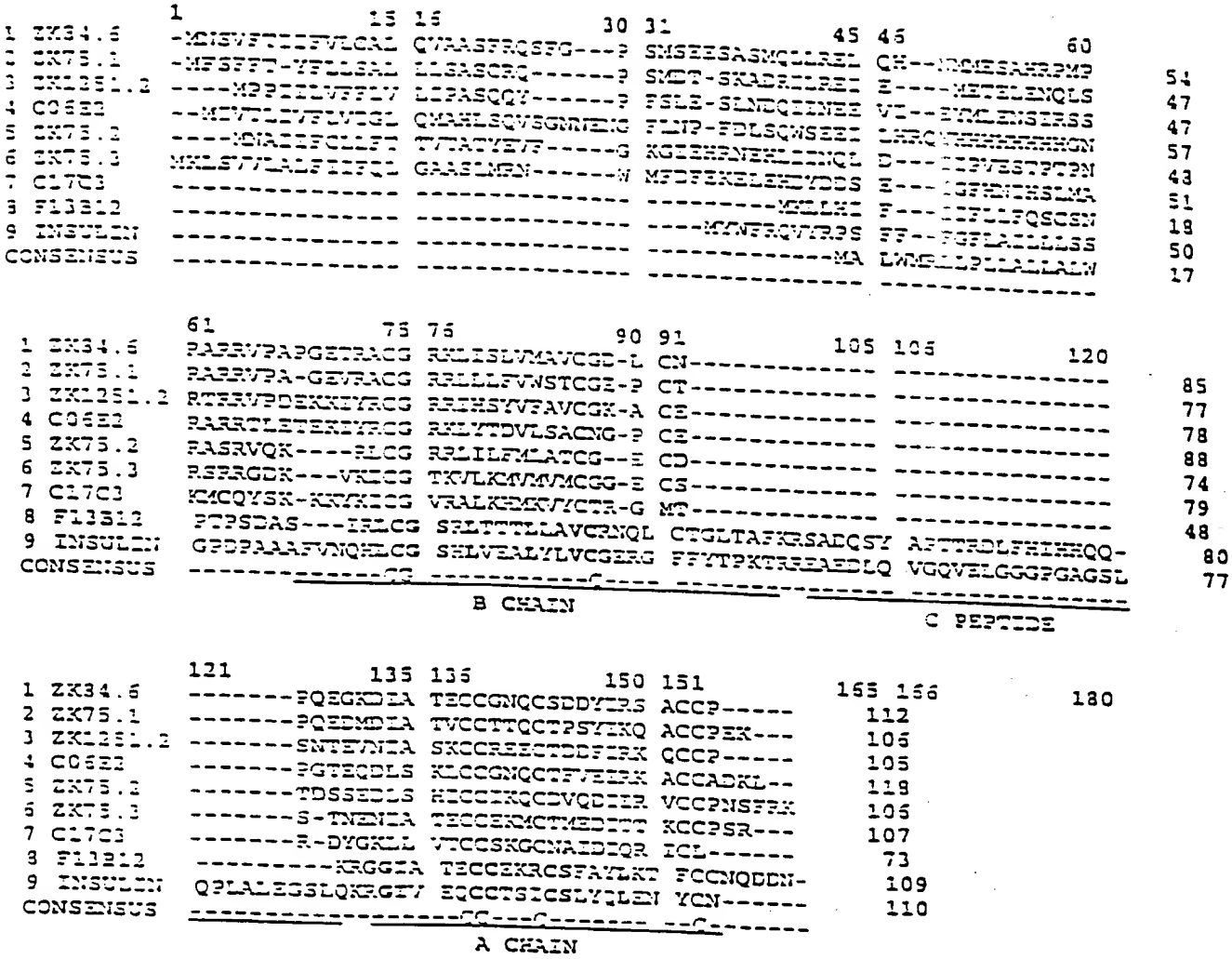


FIGURE 28

46/53

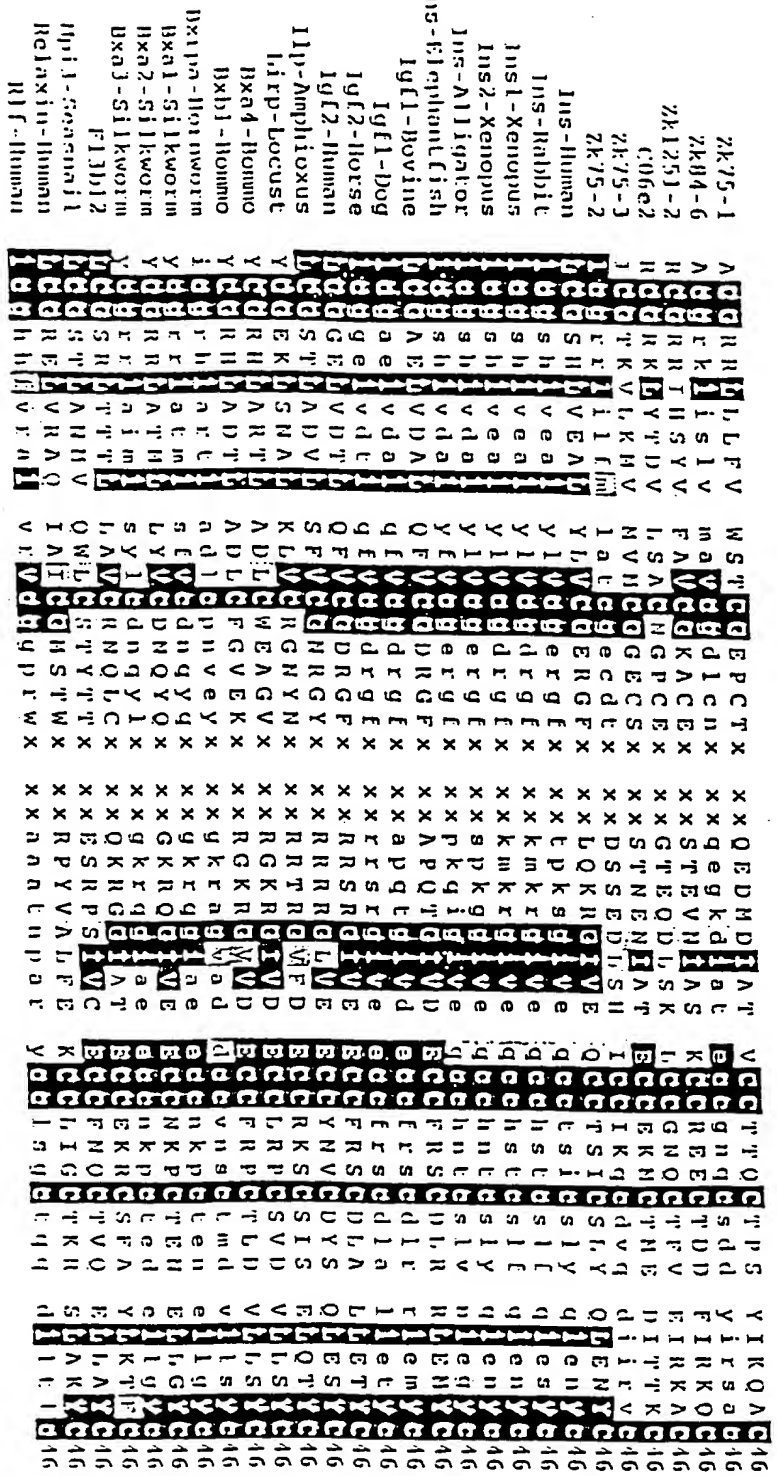


FIGURE 29

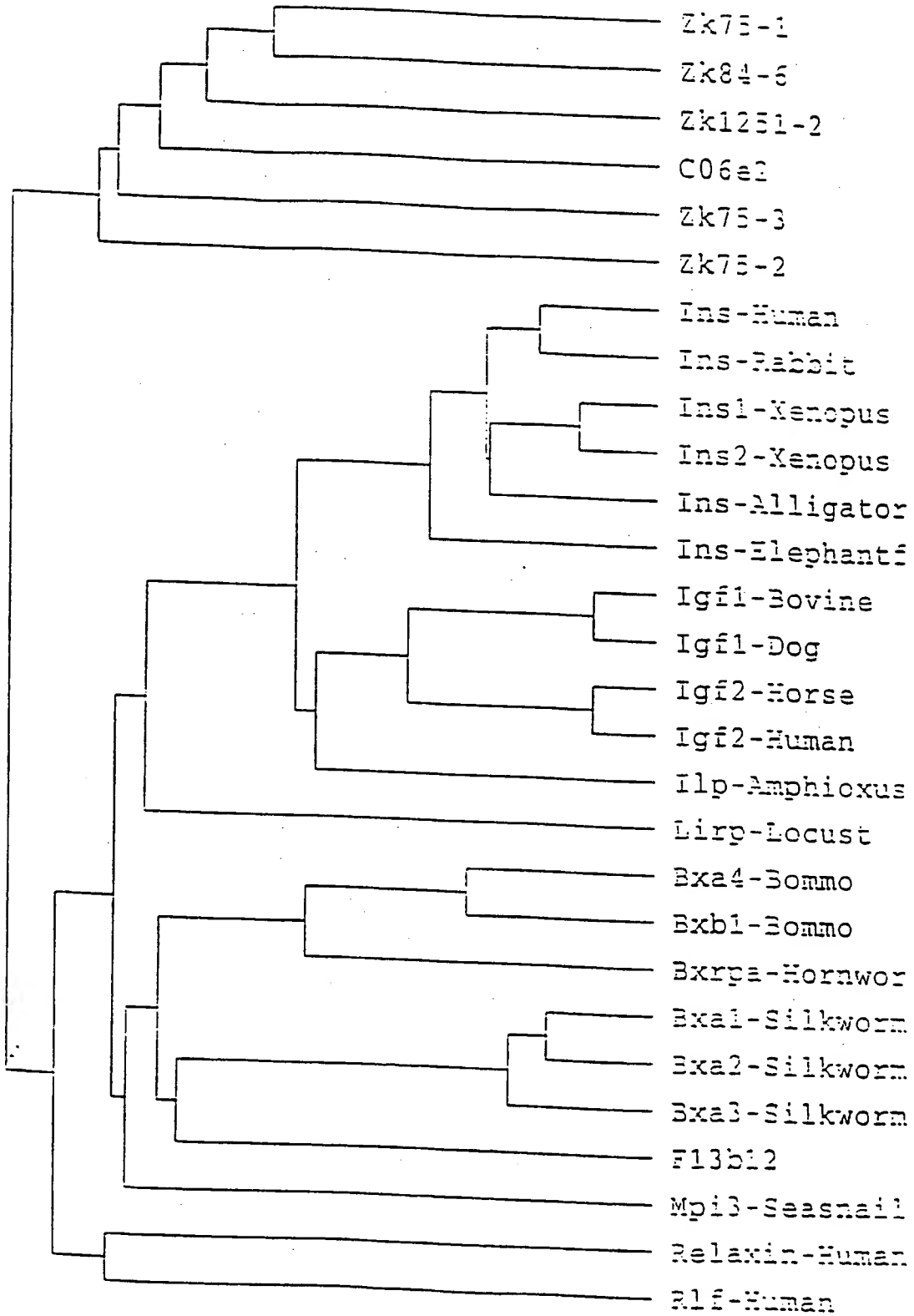

Zk75-1
Zk84-6
Zk1251-2
C06e2
Zk75-3
Zk75-2
Ins-Human
Ins-Rabbit
Ins1-Xenopus
Ins2-Xenopus
Ins-Alligator
Ins-Elephantf
Igf1-Bovine
Igf1-Dog
Igf2-Horse
Igf2-Human
Ilp-Amphioxus
Lirp-Locust
Bxa4-Bommo
Bxb1-Bommo
Bxrpa-Hornwor
Bxa1-Silkworm
Bxa2-Silkworm
Bxa3-Silkworm
F13b12
Mpi3-Seasnail
Relaxin-Human
Rlf-Human

FIGURE 30

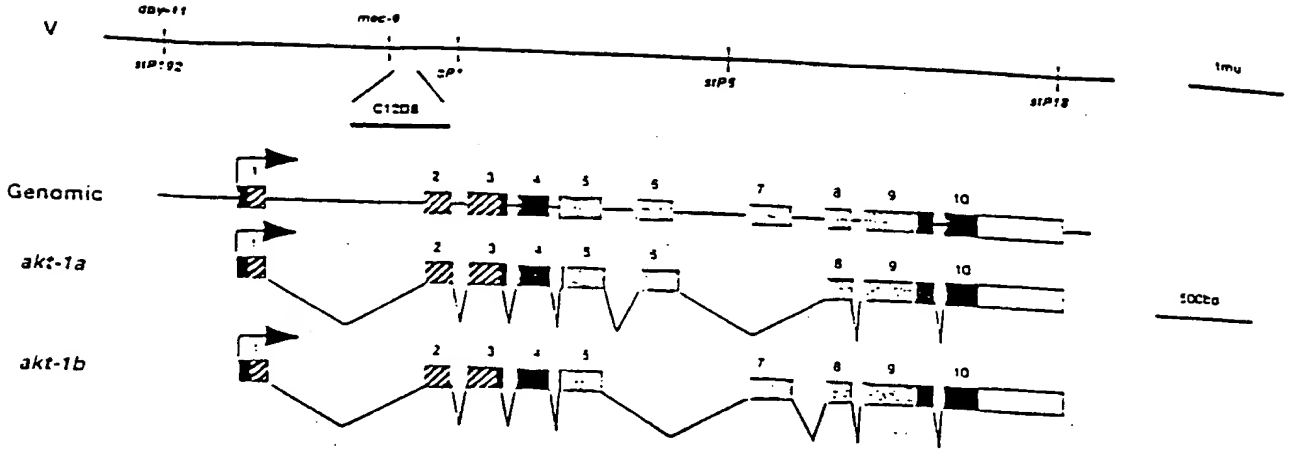


FIGURE 31

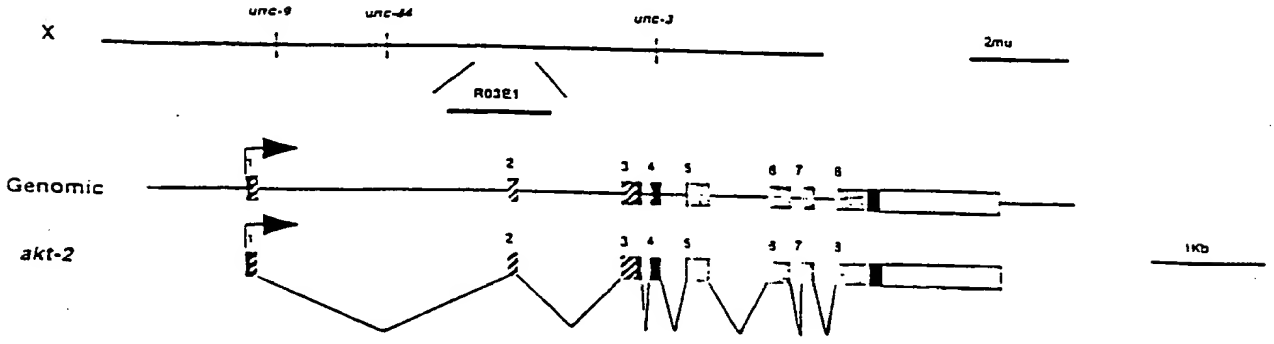


FIGURE 32

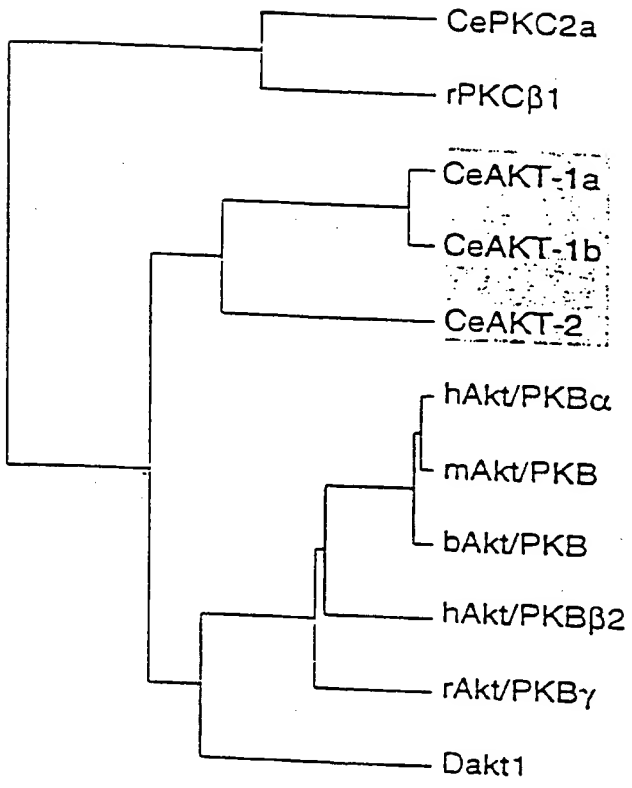

CePKC2a
rPKCβ1
CeAKT-1a
CeAKT-1b
CeAKT-2
hAkt/PKBα
mAkt/PKB
bAkt/PKB
hAkt/PKBβ2
rAkt/PKB-γ
Dakt1

FIGURE 33

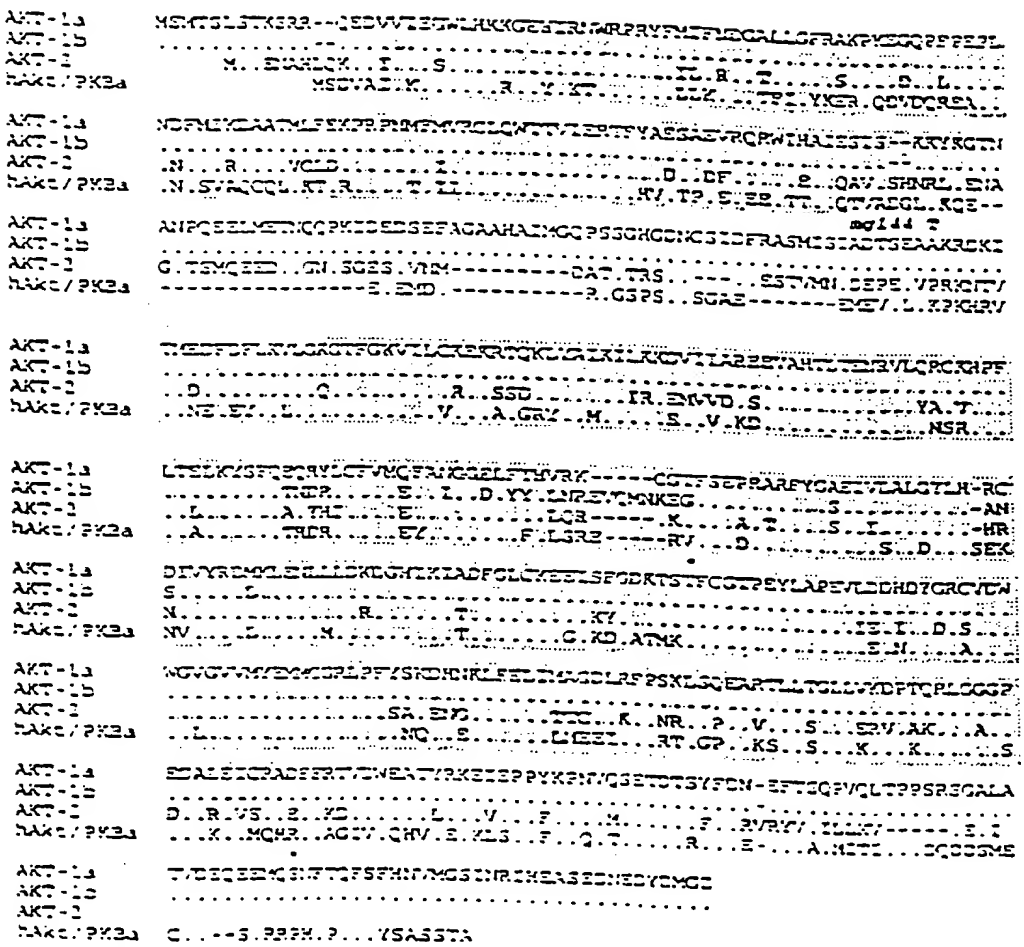


FIGURE 34

FIGURE 35A

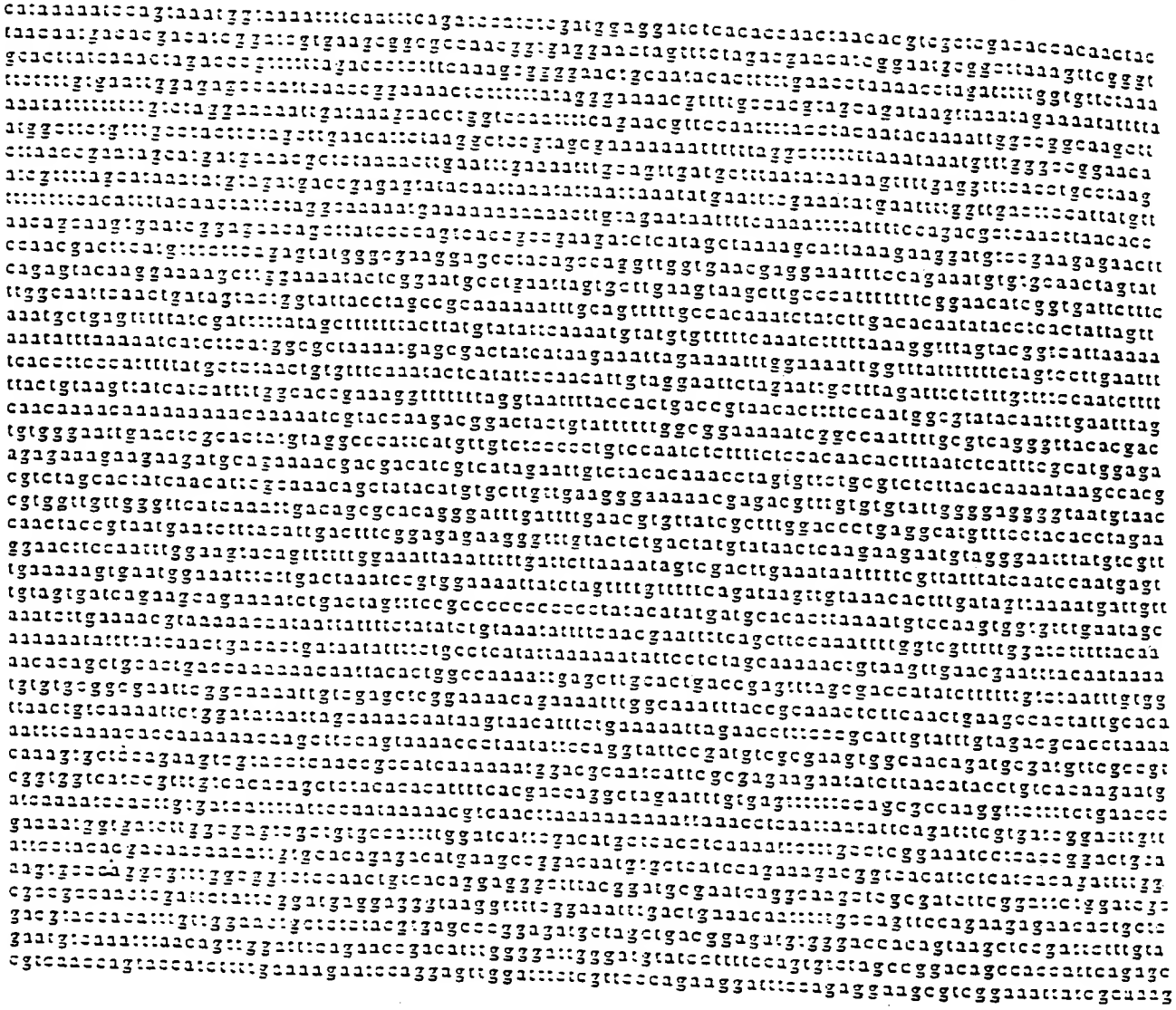
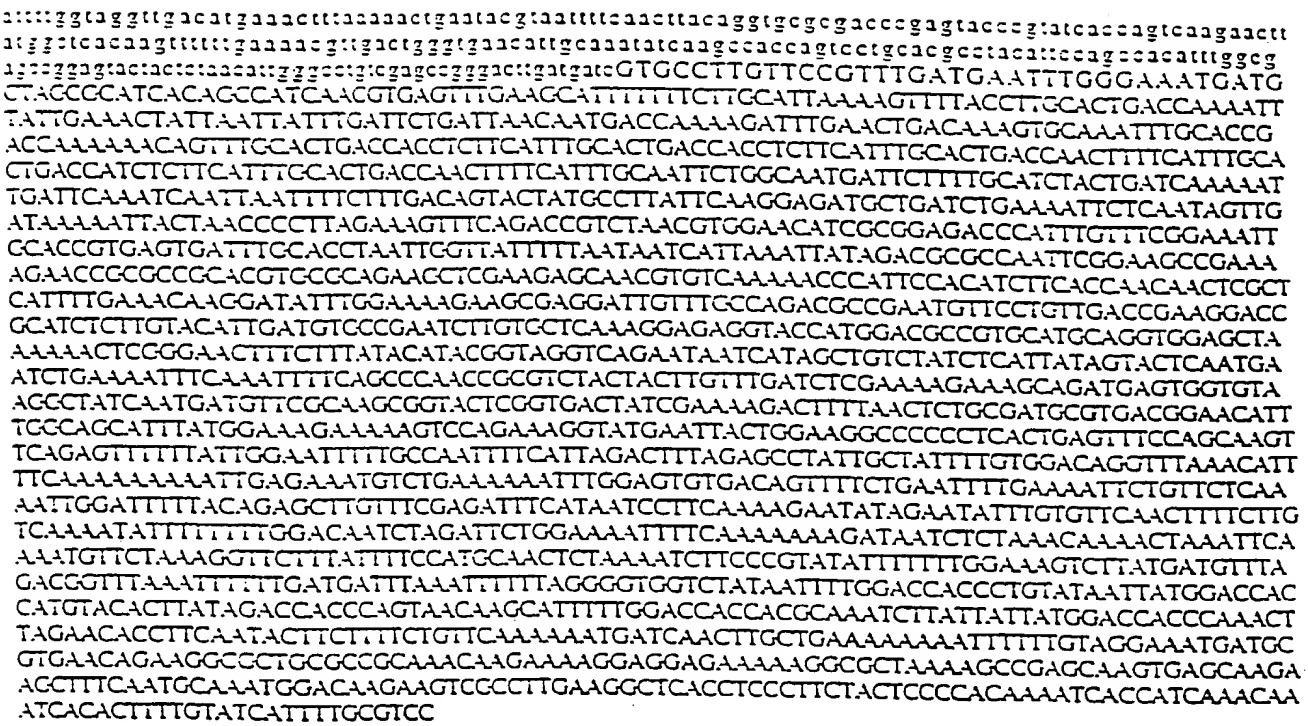


FIGURE 35B


atttggtaggttgacatgaaacttaaaaactgaatacgtaattttcaacttacaggtgcgcgacccgagtacccgtatcaccagtcaagaacttatggctcacaagtttttcgaaaacgttgactgggtgaacattgcaaatatcaagccaccagtcctgcacgcctacattccagccacatttggcgatggagtactactctaacattgggcctgtcgagccgggacttgatgatcGTGCCTTGTTCCGTTTGATGAATTTGGGAAATGATGCTAGCCGCATCACAGCCATCAACGTGAGTTTGAAGCATTTTTTTCTTGCATTAAAAGTTTTACCTTGCACTGACCAAAATTTATTGAAACTATTAATTATTTGATTCTGATTAACAATGACCAAAAGATTTGAACTGACAAAGTGCAAATTTGCACCGACCAAAAAAACAGTTTGCACTGACCACCTCTTCATTTGCACTGACCACCTCTTCATTTGCACTGACCAACTTTTCATTTGCACTGACCATCTCTTCATTTGCACTGACCAACTTTTCATTTGCAATTCTGGCAATGATTCTTTTGCATCTACTGATCAAAAATTGATTCAAATCAATTAATTTTCTTTGACAGTACTATGCCTTATTCAAGGAGATGCTGATCTGAAAATTCTCAATAGTTGATAAAAATTACTAACCCCTTAGAAAGTTTCAGACCGTCTAACGTGGAACATCGCGGAGACCCATTTGTTTCGGAAATTGCACCGTGAGTGATTTGCACCTAATTGGTTATTTTTAATAATCATTAAATTATAGACGCGCCAATTCGGAAGCCGAAAAGAACCGCGCCGCACGTGCCGCAGAACGTCGAAGAGCAACGTGTCAAAAACCCATTCCACATCTTCACCAACAACTCGCTCATTTTGAAACAAGGATATTTGGAAAAGAAGCGAGGATTGTTTGCCAGACGCCGAATGTTCCTGTTGACCGAAGGACCGCATCTCTTGTACATTGATGTGCCCGAATCTTGTGCTCAAAGGAGAGGTACCATGGACGCCGTGCATGCAGGTGGAGCTAAAAAACTCCGGGAACTTTCTTTATACATACGGTAGGTCAGAATAATCATAGCTGTCTATCTCATTATAGTACTCAATGAATCTGAAAATTTCAAATTTTCAGCCCAACCGCGTCTACTACTTGTTTGATCTCGAAAAGAAAGCAGATGAGTGGTGTAAGCCTATCAATGATGTTCGCAAGCGGTACTCGGTGACTATCGAAAAGACTTTTAACTCTGCGATGCGTGACGGAACATTTGCCAGCATTTATGGAAAGAAAAAGTCCAGAAAGGTATGAATTACTGGAAGGCCCCCCTCACTGAGTTTCCAGCAAGTTCAGAGTTTTTTATTGGAATTTTTGCCAATTTTCATTAGACTTTAGAGCCTATTGCTATTTTGTGGACAGGTTTAAACATTTTCAAAAAAAAAATTGAGAAATGTCTGAAAAAATTTGGAGTGTGACAGTTTTCTGAATTTTGAAAATTCTGTTCTCAAAATTGGATTTTTACAGAGCTTGTTTCGAGATTTCATAATCCTTCAAAAGAATATAGAATATTTGTGTTCAACTTTTCTTGTCAAAATATTTTTTTTGGACAATCTAGATTCTGGAAAATTTTCAAAAAAAGATAATCTCTAAACAAAACTAAATTCAAAATGTTCTAAAGGTTCTTTATTTTCCATGCAACTCTAAAATCTTCCCGTATATTTTTTTGGAAAGTCTTATGATGTTTAGACGGTTTAAATTTTTTGATGATTTAAATTTTTTAGGGGTGGTCTATAATTTTGGACCACCCTGTATAATTATGGACCACCATGTACACTTATAGACCACCCAGTAACAAGCATTTTTGGACCACCACGCAAATCTTATTATTATGGACCACCCAAACTTAGAACACCTTCAATACTTCTTTTCTGTTCAAAAAATGATCAACTTGCTGAAAAAAAATTTTTTGTAGGAAATGATGCGTGAACAGAAGGCGCTGCGCCGCAAACAAGAAAAGGAGGAGAAAAAGGCGCTAAAAGCCGAGCAAGTGAGCAAGAAGCTTTCAATGCAAATGGACAAGAAGTCGCCTTGAAGGCTCACCTCCCTTCTACTCCCCACAAAATCACCATCAAACAAATCACACTTTTGTATCATTTTGCGTCC

MEDLTPTNTSLDTTTTNNDTTSDREAAPTTLNLTPTASESENSLSPVTAEDLIAKSIKEGCPKRTSNDFMFLQSMGEG
AYSQVFRCREVATDAMFAVKVLQKSYLNRHQKMDAIREKNILTYLSQECGGHPFVTQLYTHFHDQARIYFVIGLV
ENGDLGESLCHFGSFDMLTSKFFASEILTGLQFLHDNKIVHRDMKPDNVLIQKDGHILITDFGSAQAFGGLQLSQEGFT
DANQASSRSSDSGSPPPTRFYSDEEEENTARRTTFVGTALYVSPEMLADGDVGPQTDIWGLGCILFQCLAGQPPFRAV
NQYHLLKRIQELDFSFPEGFPEEASEILAKILVRDPSTRITSQELMAHKFFENVDWVVNIANIKPPVLHAYIPATFGEP
EYYSNIGPVEPGLDDRALFRLMNLGNDASASQPSTFRPSNVEHRGDPFVSELAPRANSEAEKNRAARAQKLEEQRVK
NPFHIFTNNSLILKQGYLEKKRGLFARRRMFLLTEGPHLLYIDVPNLVLKGEVPWTPCMQVELKNSGTFFIHTPNR
VYYLFDLEKKADEWCKAINDVRKRYSVTIEKTFNSAMRDGTFGSIYGKKKSRKEMMREQKALRRKQEKEEKKAL
KAEQVSKKLSMQMDKKSP

FIGURE 36

MEDLTPTNTSLDTTTTNNDTTSDREAAPTTLNLTPTASESENSLSPVTAEDLIAKSIKEGCPKRTSNDFMFLQSMGEG
AYSQVFRCREVATDAMFAVKVLQKSYLNRHQKMDAIREKNILTYLSQECGGHPFVTQLYTHFHDQARIYFVIGLV
ENGDLGESLCHFGSFDMLTSKFFASEILTGLQFLHDNKIVHRDMKPDNVLIQKDGHILITDFGSAQAFGGLQLSQEGFT
DANQASSRSSDSGSPPPTRFYSDEEVPEENTARRTTFVGTALYVSPEMLADGDVGPQTDIWGLGCILFQCLAGQPPFR
AVNQYHLLKRIQELDFSFPEGFPEEASEILAKILVRDPSTRITSQELMAHKFFENVDWVVNIANIKPPVLHAYIPATF
GEPEYYSNIGPVEPGLDDRALFRLMNLGNDASASQPSTFRPSNVEHRGDPFVSELAPRANSEAEKNRAARAQKLEE
QRVKNPFHIFTNNSLILKQGYLEKKRGLFARRRMFLLTEGPHLLYIDVPNLVLKGEVPWTPCMQVELKNSGTFFIH
TPNRVYYLFDLEKKADEWCKAINDVRKRYSVTIEKTFNSAMRDGTFGSIYGKKKSRKEMMREQKALRRKQEKEE
KKALKAEQVSKKLSMQMDKKSP

FIGURE 37

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/10080

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) : A61K 49/00; C12N 5/06; C07H 21/04
US CL :424/9.1, 9.2; 800/2; 536/23.5
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : 424/9.1, 9.2; 800/2; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ESTEVEZ et al. The daf-4 gene encodes a bone morophogenetic protein receptor controlling C. elegans dauer larva development. Nature. 14 October 1993, Vol. 365, pages 644-649, especially figure 1.	1-3, 7, 8, 10, 11, 18, and 22
X -- Y	GALILI et al. Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosacoma. Nature Gentics. November 1993. Vol. 5, No. 3, pages 230-235 (erata sheet attached, Nature Genetics, Vol. 6, No. 2, page 214-214). see entire document.	10 --- 1-3, 7, 8, 10, 11, 18, and 22

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"B"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
18 JULY 1998

Date of mailing of the international search report
28 SEP 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
CHRISTOPHER S. F. LOW
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/10080

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	ZWALL et al. Target-selected gene inactivation in Caenorhabditis elegans by using a frozen transposon insertion mutant bank. Proc. Natl. Acad. Sci. USA. August 1993, Vol 90, pages 7431-7435, see entire document.	10 ---------- 1-3, 7, 8, 10, 11, 18, and 22
Y	MURAKAMI et al. A Genetic Pathway Conferring Life Extension and Resistance to UV Stress in Caenorhabditis elegans. Genetics. July 1996, Vol. 143, pages 1207-1218, see entire document.	1-3, 7, 8, 10, 11, 18, and 22
X,P ---- Y,P	LIN et al. daf-16: An HNF-3/forkhead Family Member Than Can Function to Double the Life-Span of Caenorhabditis elegans. Science. 14 November 1997, Vol. 278, pages 1319-1322, see entire document.	10 ---------- 1-3, 7, 8, 10, 11, 18, and 22
X,P ----- Y,P	OGG et al. The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. Nature. 30 October 1997, Vol. 389, pages 994-999, see entire document.	10 ---------- 1-3, 7, 8, 10, 11, 18, and 22
X -- Y	MCCOMBIE et al. Caenorhabditis elegans expressed sequence tags identify gene families and potential disease gene homologues. Nature Genetics. May 1992, Vol. 1, pages 124-131, see entire document.	10 ---------- 1-3, 7, 8, 10, 11, 18, and 22
X -- Y	WATERSTON et al. A survey of expressed genes in Caenorhabditis elegans. Nature Genetics. May 1992, Vol. 1, pages 114-123, see entire document.	10 ---------- 1-3, 7, 8, 10, 11, 18, and 22
Y	KENYON et al. A C. elegans mutant that lives twice as long as wild type. Nature. 02 December 1993, Vol. 366, 461-464, see entire document.	1-3, 7, 8, 10, 11, 18, and 22
Y	REN et al. Control of C. elegans Larval Development by Neuronal Expression of a TGF-Beth Homolog. Science. 22 November 1996, Vol. 274, pages 1389-1391, see entire document.	1-3, 7, 8, 10, 11, 18, and 22
Y, P	KIMURA et al. daf-2, an Insulin Receptor-Like Gene That Regulates Longevity and Diapause in Caenorhabditis elegans. Science. 15 August 1997, Vol. 277, pages 942-946. see entire document.	1-3, 7, 8, 10, 11, 18, and 22
Y	US 5,196,333 A (CHALFIE et al.) 23 March 1993, see entire document.	1-3, 7, 8, 10, 11, 18, and 22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/10080

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3, 7, 8, 10, 11, and 22

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/10080

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

Automated Patent System - files USPAT, EPOABS, JPOABS
DIALOG ONESEARCH - BIOSIS
Search terms: nematode, Caenorhabditis elegans, transgene, transgenic, decay, activating, factor, DAF-2, DAF-3, DAF-7, DAF-16, DAF, AGE-1

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

I. Claim 1-3, 7, 8, 10, 11, 18, 22 drawn to a first method of in vivo testing to detect compounds that decrease activity of DAF using a transgenic animal which produces any heterologous decay activation factor (DAF) are, classified in Class 424, subclass 9.2; a product (claim 10, the DNA encoding the DAF16 polypeptide); and a first method of making (claim 22) a product of the first method of use.

II. Claim 4, drawn to a method of testing in vitro using cells which produce heterologous decay activation factor (DAF).

To the extent that claims 5 and 6 depend from claims 1 or 4, claims 5 and 6 will be examined with Group I and/or II should additional fees be paid for a search of Group II.

III. Claim 9, drawn to a method of in vivo testing a human gene for involvement impaired glucose tolerance or obesity using a nematode which has a mutation in a daf or an age gene and is a second method of use. The claims of Group I do not require an *age* gene nor the human gene for practice.

IV. Claims 12-14, 19-21 drawn to a third process of use directed to isolation of a gene or a portion thereof with SEQ ID Nos: 54 through 57 by hybridization.

V. Claims 15-17, drawn to a treatment modality that delays onset of impaired glucose tolerance or obesity by administering a compound that inhibits DAF-16 or DAF-3 polypeptide.

VI. Claim 23, drawn to a method of diagnosis of impaired glucose tolerance or obesity by identification of DNA that encodes a mutation in a DAF gene.

The inventions listed as Groups I through VI do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons.

The present written description (page 21) defines the DAF-16 polypeptide as at least 71 or 35 or 65 or 53% identical to SEQ ID Nos 54 or 55 or 56 or 57 respectively. Claim 10 of Group I, the line product calm is interpreted as a DNA encoding the above defined protein.

The Galili *et al.* (1993) Nature Genetics 5: 230-235 reference disclosed isolated genetic material encoding a fork head domain protein, which contains, absent factual data to the contrary, the genetic material encoding a *C. elegans* DAF-16 polypeptide. In view of the foregoing, the claims of Group I in regard to the first claimed product, the special technical feature of which as presently claimed does not define a contribution considered as a whole over the prior art.

The claim of Group II (an *in vitro* process) is directed to a different process from that of Group I (an *in vivo* process); and, Groups III through VI are directed to alternative processes of use.